The Problem of Data Adequacy in Applied Statistics

Von der Mercator School of Management

– Fakultät für Betriebswirtschaftslehre –

der Universität Duisburg-Essen (Campus Duisburg)

zur Erlangung des akademischen Grades

eines Doktors der Wirtschaftswissenschaft (Dr. rer. oec.)

genehmigte Dissertation

von

Jens Mehrhoff

aus

Minden

Referent:	Professor Dr. Peter Anker
Korreferent:	Professor Dr. Jens Südekum
Tag der mündlichen Prüfung:	2. Juni 2010

Für meine Eltern,
die mich auf meinem Weg immer unterstützt haben.

Danksagung

Mein besonderer Dank gilt meinem Doktorvater und dem Betreuer meiner Doktorarbeit, Professor Dr. Peter von der Lippe und Professor Dr. Peter Anker, ohne deren tatkräftige Unterstützung bei meiner Promotion diese nicht möglich gewesen wäre. In ausführlichen Diskussionen konnte ich von der Erfahrung der beiden profitieren und erhielt zahlreiche förderliche Anregungen für die Verbesserung meiner Arbeit. Meine Dissertation hätte sicherlich nicht so reibungslos zum erfolgreichen Abschluss gebracht werden können, wenn ich nicht durch die fachliche Betreuung und den persönlichen Einsatz von ihnen allzeit gefördert worden wäre. Professor Dr. Jens Südekum und Professorin Dr. Antje Mahayni danke ich ebenfalls für ihre äußerst konstruktive Mitwirkung in meinem Promotionsverfahren.

Weiter danke ich sehr herzlich meinen Vorgesetzten bei der Deutschen Bundesbank aus dem Zentralbereich Statistik, Herrn Robert Kirchner und Herrn Hans-Albert Leifer, denen ich die Möglichkeit zur Durchführung meiner Promotion sowie den Vorschlag zur Fragestellung schulde. Die exzellente fachliche Begleitung hat erheblich zum Gelingen meiner Arbeit beigetragen. Durch viele sehr nützliche Kommentare habe ich die Qualität dieser maßgeblich verbessern können. Insbesondere die Fokussierung auf die praktische Relevanz und die strukturierte Darstellung meiner Ergebnisse waren für mich die Schlüssel zum Erfolg. Ferner empfand ich die mir stets eingeräumten Freiheiten bei der Umsetzung meiner Ideen und Verwirklichung meiner Interessen als sehr angenehm.

Schließlich gilt mein spezieller Dank Frau Dr. Sophia Müller-Spahn. Durch zahllose wertvolle Kommentare und viele anregende Diskussionen hat sie wesentlichen Einfluss auf meine Forschung genommen. Sie zeigte mir andere Sichtweisen auf und übte konstruktive Kritik, welche die Schwachstellen meiner Argumentation offenbarte.

The Problem of Data Adequacy in Applied Statistics

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

This thesis focuses on some important methodological issues in applied statistics. More precisely, the adequacy of data for empirical research is discussed. In particular, three stages of the process of producing and using official statistics are investigated and their degree of uncertainty is quantified. The stages are data processing, validation and analysis. At each of these three stages one of the following research questions is answered. First, the producer view is taken by asking "how should price indices be calculated from micro data in official statistics?" Second, the views of both users and producers are considered by posing the question "how reliable are timely published official statistics partially based on estimates?" Last, the problem "how should one estimate econometric models with micro data from official statistics?" is treated from the user perspective. In answering these questions the following significant contributions to the literature are made in each of the three main chapters.

The first main chapter analyses the processing of raw data in terms of calculating elementary price indices in foreign trade statistics. Most of the literature on elementary price indices discusses the choice of a particular index formula based on the axiomatic approach (cf. Eichhorn, 1978, and Diewert, 1995). This approach states properties which an index formula should desirably fulfil and checks which axioms are actually fulfilled. The importance of axioms in general depends heavily on the purpose of the index formula in question and, to some extent, on personal preferences. In any case, the axiomatic approach is of little guidance in choosing the elementary index (for which weights are not available) corresponding to the characteristics of the index at the second stage (where weights are actually available). It exclusively deals with the mathematical properties of an index formula. Thus, it is an all or nothing decision in favour of or against it. It consequently completely neglects the most relevant issue in practice as to what extent a condition is not fulfilled. The statistical approach newly developed in this chapter fills this void. It contributes to the literature by looking at how numerical equivalence between an unweighted elementary index and a weighted aggregate index can be achieved, independent of the axiomatic properties. It is shown that the solution to the problem of elementary indices that correspond to a desired aggregate index depends on the empirical correlation between prices and quantities, in particular on the price elasticity. Based on this, consistency between price and volume measurement is achieved. In addition to the analytical derivation, this is demonstrated empirically in an application using data from German foreign trade statistics.

In the second main chapter the validity of the data is checked by decomposition of revisions of real time data in a seasonal adjustment context. Revisions occur if preliminary data are updated and estimates are replaced with actual figures. Moreover, new data lead to revisions of the seasonally adjusted time series – even if old preliminary data remain unchanged. The limited literature on decomposition of revisions deals almost completely with the question of whether revisions are "news or noise" (cf. Mankiw and Shapiro, 1986). This strand of the literature discusses the informational content of revisions. Revisions are considered to be "news" if they are orthogonal to the full information set available at the time the first estimate was published, i.e. they are unpredictable. Vice versa, if revisions are indeed predictable, the first estimate is a "noisy" measure of the most recent one, hence it is an inefficient forecast as revisions correlate with a subset of the data available for the first estimate. In contrast, this chapter is concerned with decomposition of revisions into their sources. For that, a new procedure is developed and implemented within the framework of the seasonal adjustment method X-12-ARIMA which is the contribution to the literature. This is relevant because the ability of a seasonal adjustment method to produce low revisions from its technical procedure can be thought of as being a quality characteristic for it. In an empirical application to five important German business cycle indicators, revisions of unadjusted real time data are found to play a larger role than those stemming from the seasonal adjustment method. This result is not self-evident as, for example, for European time series it is frequently argued that the seasonal adjustment method is the main reason for revisions.

The focus of the third main chapter is on data analysis; here, dynamic panel data models are estimated by means of GMM with more robust, in this case factorised, instruments. Almost all of the empirical literature and the major proportion of the theoretical literature focus exclusively on the exogeneity assumption. Exogeneity means that the instruments are uncorrelated with the error term. If this were true, white noise processes would be ideal instruments as they per definitionem correlate with nothing. But in fact this is only half of the story: white noise processes as instruments are consistent with virtually any estimate of the parameters of interest. The other half of the story is the relevance assumption (cf. Staiger and Stock, 1997). This assumption states that the instruments correlate with the endogenous regressor(s). The weaker the instrument set becomes, the worse are the small-sample properties of any instrumental variables estimator, such as GMM. The contribution of this chapter to the literature is the proposition of a methodology that has improved finite-sample properties. In particular, the instrument set is factorised so that its informational content is condensed in a much lower number of instruments employed in the estimation. This results in lower biases and lower RMSEs which is shown by Monte Carlo simulations.

1.2 Overview

The first main chapter proposes a theoretical framework which allows the achievement of numerical equivalence of an elementary index with the Laspeyres, Paasche or Fisher price index. An application using data from German foreign trade statistics illustrates the methodology outlined here.

It is customary in official statistics, although often neglected in theoretical papers, for most price indices to be calculated in two stages. At the first stage, elementary indices are calculated on the basis of prices or their relatives, without having information on quantities or expenditures. At the second stage, the aggregate index is calculated on the basis of the elementary indices from the first stage, using aggregate expenditure share weights.

In general, the question of "what should be measured?" directly yields the optimal index formula at the second stage – in this instance, the Laspeyres, Paasche or Fisher price index. The Laspeyres price index ensures the principle of pure price comparison over multiple periods by using a fixed basket of goods and is consistent in aggregation. The Paasche price index leads to volume measures in constant prices which are consistent in aggregation and purely comparable over multiple periods (Laspeyres principle for the quantity index). Among other so-called cost of living indices, the Fisher price index approximates the change in the minimum expenditures, which preserve utility at a constant level, owing to changes in (relative) prices.

However, it needs to be determined which index formula at the elementary level, where no expenditure share weights are available, corresponds to a desired aggregate index. Due to the unavailability of these weights for every single good or service, an equally weighted index formula is used at the first stage. The importance of the elementary level and the elementary index cannot be emphasised enough. Biases of these indices at this level are more severe than the pros and cons of the formula at the aggregate level. This is because the two-staged index can never be better than its building blocks. No matter how good the aggregate expenditure share weights are, they simply cannot compensate biases of the elementary indices – in fact, they just weight them together.

The existing approaches to index numbers including but not restricted to the axiomatic approach are of little guidance in choosing the elementary index corresponding to the characteristics of the index at the second stage. In order to achieve numerical equivalence between an elementary index and an arbitrary aggregate index, a statistical approach is developed. The basic idea behind this approach is that different elementary indices implicitly weight price relatives differently, although they do not imply an explicit expenditure structure.

It is firstly demonstrated that every weighted index can be expressed one-to-one and onto as a "power mean". The power cannot be derived analytically without making further assumptions. Here, the solution to the problem of corresponding elementary indices depends on a parametric joint distribution of prices and quantities and their functional relation. Hence, secondly, the log-normal distribution is introduced. A closed form solution is provided as to which power corresponds to a given aggregate index. Thirdly, the log-normal distribution parameters are related to the price elasticity. It is assumed that an equilibrium quantity traded for each good and time exists, and that the adjustment to this equilibrium is both incomplete ("partial adjustment") and erroneous. Finally, it is shown that the choice of the elementary indices which correspond to the desired aggregate ones can be based on the price elasticity alone. A power mean with power equal to minus the price elasticity yields approximately the same result as the Laspeyres price index; however, if the Paasche price index should be replicated, the power of the power mean must equal the price elasticity. A quadratic mean of order two times the absolute price elasticity corresponds to the Fisher price index.

This is also demonstrated empirically in an application using data from German foreign trade statistics. The data set covers 1,264 panels consisting of 12,948 goods, for exports as well as for imports, and a total of 1,839,384 observations over the period January 2000 to December 2007. Panel unit root tests show stationarity of both prices and quantities for almost all panels in exports as well as in imports. The price elasticity is estimated in the framework of a log-linear partial adjustment model by means of dynamic panel data one-step system GMM.

For the Laspeyres price index as desired aggregate index, 70% of the groups of goods in exports and 72% in imports imply the use of the Carli index at the elementary level. Regarding trade values these figures reduce to 62% and 66%, respectively. The Jevons index performs best at the first stage in 14% of the groups in exports and 17% in imports with much higher shares with respect to trade value, 29% and 28%, respectively. In 15% of the groups in exports and 10% in imports, the quadratic index is desirable at the lower level of aggregation, trade value shares here are 7% and 5%, respectively. Shares missing to 100% reflect other indices. If the Paasche price index is taken as desired aggregate index, the corresponding power means are inverted: instead of the Carli index the harmonic index, and instead of the quadratic index the reciprocal quadratic index have to be used; if the Jevons index corresponds to the Laspeyres price index, it does so for the Paasche price index, too. If the desired aggregate index is chosen to be the Fisher price index, the results are as follows. Jevons index: 6% in exports (trade value: 20%) and 7% in imports (17%); Hybrid index: 21% (19%) and 28% (25%); CSWD index: 46% (48%) and 44% (43%), cubic order: 21% (9%) and 15% (12%); quartic order: 6% (3%) and 4% (2%). Again, quintic and higher orders make up shares missing to 100%.

The conclusions are twofold. Firstly, index calculation can be rendered more precise if different elementary indices are applied to each group of goods, reflecting their specific price elasticities. This has to be done in order to come as close as possible to the optimal aggregate Laspeyres, Paasche or Fisher price index. Secondly, for different purposes – either price or volume measurement – different elementary indices should be calculated for the same data. This means that if the Carli index is applied as the single formula at the elementary level of a Laspeyres price index, implying a price elasticity of minus one, the harmonic index must be used at the elementary level of a Paasche price index.

The second main chapter deals with revisions to seasonally adjusted real time data. The sources for these revisions are, on the one hand, those from the seasonal adjustment method and, on the other hand, those from unadjusted real time data. The revisions are empirically quantified and a decomposition procedure is presented as to how much each of the sources contributes to total revisions.

The importance of real time data becomes obvious when one tries to understand economic policy decisions made based on historical data and reconsiders these past situations in the light of more recent data. Statistical agencies and users of seasonally adjusted real time data alike are interested in it, inter alia in terms of the quality and interpretation of statistics. Thus, revisions of real time data are a frequently discussed topic. Revisions to seasonally adjusted real time data have two separate but inter-related sources. One of these sources is the technical procedure of the method used for seasonal adjustment, while the other is the revisions process of unadjusted data in real time. These two sources are empirically quantified and total revisions are decomposed into them.

The decomposition of a time series into unobservable components forms the basis of seasonal adjustment. It is assumed that the unadjusted time series is decomposed into a trend-cycle component, a seasonal component (assume, for the sake of simplicity, that calendar effects are included in the seasonal component) and an irregular component. Seasonal adjustment is based on the Census X-12-ARIMA method – an iterative, mathematical procedure for seasonally adjusting time series.

The purpose of seasonal adjustment is to filter out the usual seasonal fluctuations, i.e. those movements that recur with similar intensity in the same season each year. Hence, this implies that the seasonally adjusted series will still display fluctuations due to exceptionally strong or weak seasonal influences. Besides this, unusual movements that are readily understandable in economic terms as well as other random disruptions will also continue to be visible. However, seasonally adjusting includes the elimination of working-day variations insofar as influences that derive from differences in the number of working days or the dates of particular days can be demonstrated and quantified.

These regular patterns contain no news for analysing the current economic development. Thus, the intention of seasonal adjustment is to eliminate these patterns from the unadjusted time series, whilst maintaining news in the seasonally adjusted time series – smoothness of this series, on the contrary, is no criterion. Eventually, period-to-period changes of the seasonally adjusted time series are to be interpreted economically; however, their movements are at least partly random noise.

Revisions of the seasonally adjusted time series are defined as the deviation of the most recent estimate from the first one (for the same reporting date). This can be interpreted as the answer to the question of how much a given adjustment is affected by appending new and updating old unadjusted data. While the former is based on all unadjusted data available at the most recent date of release, the latter is based only on unadjusted data up to the first publication of that reporting date. Hence, a trade-off between timeliness and accuracy arises.

Data used in this study are firstly unadjusted real time data rebased to the current base year, and secondly the present user setting of seasonal adjustment which is held constant throughout all vintages. For this analysis a new procedure is developed and implemented within the framework of the seasonal adjustment method X-12-ARIMA. Using the two aforementioned data sources, the real time data is seasonally readjusted, i.e. historical published figures are not used. The period covered is from the beginning of 1991 to the end of 2006. The analysis of revisions is based on the six-year period from 1996 to 2001.

As revisions of unadjusted data influence the estimation of the seasonal component, a variance decomposition approach becomes necessary. The framework for decomposition of revisions is a heterogeneous panel regression model. As estimated slope coefficients indicate marginal effects, elasticities are calculated and set in relation to each other. The investigated time series are important business cycle indicators for Germany: (1) real gross domestic product, (2) employment, (3) output in and (4) orders received by the manufacturing sector as well as (5) retail trade turnover.

It can be concluded that revisions of unadjusted real time data play a larger role when explaining revisions of seasonally adjusted real time data for Germany as their elasticities were greater than those of seasonal adjustment. Since these imply the use of newly available information, compared to technical revisions from the seasonal adjustment method, they are the lesser of two evils. Furthermore, this analysis confirmed a well-known result for the recent past: the current domain of uncertainty of seasonal adjustment depends heavily on the time series analysed and their properties.

The well-known problem of too many instruments in dynamic panel data GMM is dealt with in detail in the third main chapter. It goes one step further by providing a solution to this problem: factorisation of the standard instrument set is shown to be a valid transformation for ensuring consistency of GMM. The researcher's choice of a particular transformation can be replaced by a data-driven statistical decision.

Dynamic panel data (DPD) models have become increasingly popular. They are characterised by two features. The first one is their dynamic structure, i.e. the model equation has at least one lagged dependent variable on the right-hand side. The second one is their panel structure, i.e. the data have both a cross-sectional and a time series dimension. Nowadays micro level data, such as of firms or banks, enable researchers to identify economic relationships at a disaggregate level. Examples include financing constraints of firms and interest rate pass-through of banks. Not only individual effects can be estimated with the aid of panel data – also the problem of aggregation bias can be avoided, where an aggregate regression is said to suffer from aggregation bias when the aggregate regression slope parameter does not correctly reflect the average of the individual slope parameters. However, DPD models are a source for biases themselves. The Least Squares Dummy Variables (LSDV) estimator has a non-vanishing bias for small T and large N, in particular it is downward biased.

In order to see why this is the case, assume that a cross-section faces a positive shock in one period which is not modelled and thus appears in the error term. Ceteris paribus, the associated fixed effect, which measures the difference between the average of the dependent variable that remains unexplained by other regressors and the sample average, will be higher for the whole sample period.

In the period following the shock, both the lagged dependent variable and the fixed effect will be higher. Hence, a regressor and the error term are positively correlated, violating the necessary assumption for consistency of Ordinary Least Squares (OLS). Especially, predictive power is assigned to the coefficient estimate for the lagged dependent variable that actually should be attributed to the fixed effect. Thereby, the former is inflated and the coefficient estimate is biased upward. Under the Within Groups transformation, the lagged dependent variable correlates negatively with the transformed error term so that the LSDV estimator is downward biased.

The problem of DPD bias was solved with unbiased DPD estimators based on Generalised Method of Moments (GMM) in the 1990s: first with Difference GMM and later with System GMM. The basic idea of these estimators is that lagged levels (Difference GMM) and additionally lagged differences (System GMM) are valid instruments for the lagged endogenous variable, i.e. are uncorrelated with the transformed error term. Unlike OLS, GMM does not minimise the sum of squared errors but chooses coefficients on the regressors that satisfy all moment conditions (the moments of the errors with the instruments are zero) as well as possible. However, one issue with regard to DPD GMM still remains problematic; the number of instruments grows quadratically with T. GMM becomes inconsistent as the number of instruments becomes too large. This begs the question: "what is the optimal set of instruments?" Currently, there are two techniques in use to reduce the instrument count. One of them is limiting the lag depth, the other one is "collapsing" the instrument set. These transformations are deterministic ones of the instrument matrix. Besides the fact that no widely accepted rule of thumb for the instrument count exists, by choosing one of the aforementioned approaches, the researcher decides which transformation is to be used. Yet, the question is, "can we let the data decide how the transformation matrix should look?" The answer is found here by means of principal components analysis (PCA) of the instrument set and is shown to be "yes, we can." PCA extracts the largest eigenvalues of the estimated covariance matrix of the instruments and assembles the corresponding eigenvectors in the matrix of component loadings, the transformation matrix. The resulting DPD GMM estimator is characterised by both a lower bias and a lower root mean squared error (RMSE) than the standard techniques.

The results of a Monte Carlo simulation strongly suggest the use of factorised instruments as these produce the lowest bias and RMSE. This generates a set of instruments which reduces the uncertainty in the choice of instruments. Factorised instruments do not only perform better as regards the exogeneity assumption (the risk of overfitting endogenous variables is lowered), they are also stronger in terms of the relevance assumption (the entire instrument set is not caused to be weak by many weak instruments). The latter assumption is often neglected although it is as important as the former one for the solution of the problem of too many instruments. Furthermore, there is a clear recommendation to collapse the instrument set prior to factorisation. Preferably, the lag depth is also limited. Most importantly, the bias of standard GMM increases due to instrument proliferation. The simulation further shows that LSDV should be applied only if the time dimension is much larger than 30, while pooled OLS has clearly sub-optimal properties for the estimation of DPD.

Chapter 2

Aggregate Indices and Their Corresponding Elementary Indices

2.1 Introduction

2.1.1 Motivation

It is customary in official statistics, although often neglected in theoretical papers, for most price indices to be calculated in two stages. At the first stage, elementary indices are calculated on the basis of prices or their relatives, without having information on quantities or expenditures. At the second stage, the aggregate index is calculated on the basis of the elementary indices from the first stage, using aggregate expenditure share weights.

In general, the question of "what should be measured?" directly yields the optimal index formula at the second stage: for measuring genuine price movements, a Laspeyres price index is used; for deflation purposes, a Paasche price index is preferred; and for the "cost of living", a Fisher price index, among others, is the formula of choice. However, it is less clear which index formula should be used at the first stage, where no expenditure share weights are available. The existing approaches to index numbers including but not restricted to the axiomatic approach are of little guidance in choosing the elementary index corresponding to the characteristics of the index at the second stage.

The point in question is "how can the corresponding elementary index be selected?" The answer to this question is found by the proposition of a statistical approach. A single comprehensive framework, known as "power means", unifies the aggregate and elementary levels. With the aid of this approach, theoretical conditions under which a particular index formula at the elementary level exactly equals the desired aggregate index are identified and empirically approximated.

The remainder of the chapter is organised as follows. It continues with a review of a selection of the existing literature on elementary indices. Section 2.2 introduces basic concepts and approaches in index theory along with a more thorough explanation of the problem at the elementary level. Both the theoretical foundations of power means as well as the application to the Laspeyres, Paasche and Fisher price indices and their corresponding elementary indices are presented in detail in Section 2.3. The results of an empirical application using data from German foreign trade statistics are to be found in Section 2.4. The final section concludes.

2.1.2 Literature Review

After a long period of research into aggregate formulae and an almost equally long policy debate in Europe and the US on whether the Laspeyres or Fisher formula should be used for a consumer price index (cf. Boskin et al., 1996, 1998, and Schultze and Mackie, 2002), the focus of attention has recently moved more to the question of which index formula should be used at the elementary level. Nowadays, the capabilities of modern computers and the increasing coverage of data, first and foremost, through the advent of scanner data, enables statistical offices to calculate more refined price indices even at the elementary level (cf. Silver, 1995, Silver and Webb, 2002, Feenstra and Shapiro, 2003, Diewert, 2004, and Proceedings of the Meetings of the Ottawa Group).

Diewert (2004), and Diewert and Silver (2004, 2008) devote whole chapters in the CPI, PPI and XMPI manuals to elementary indices. They deal with virtually all topics that arise around the calculation of price indices at the elementary level. Theoretical issues, such as the problem of aggregation, are covered as well as practical questions, such as numerical relationships between different elementary indices. They continue by outlining the classical approaches in index theory, i.e. the axiomatic, economic, sampling and stochastic approaches (cf. Subsection 2.2.3 for a discussion of all four approaches), and discuss the use of scanner data (cf. Subsection 2.5.2 for an outlook on a prospective study). Currently, there is an active ongoing discussion at Eurostat's Working Group on Harmonisation of Consumer Price Indices – more specifically, in the Task Force on Sampling – on which index formula is to be used at the elementary level (cf. EC, 2001, Section I). The Commission Regulation (EC, 1996, Article 7 in conjunction with Annex II) abandons the use of the Carli index but allows the use of either the Jevons or Dutot index (cf. Subsection 2.3.1.1 for the definitions of the formulae). More precisely, the Carli index is not prohibited de jure but de facto as it would have to be shown that the results do not differ by more than one-tenth of a percentage point from either the Jevons or Dutot index (cf. the next-but-one paragraph for empirical evidence and Subsection 2.3.1.2 for the mathematical relation).

Balk (1994) discusses the index formula problem at the elementary level. He poses the question whether ratios of average prices or an average of price relatives, and which type of average, i.e. arithmetic, geometric or harmonic, should be used. Turvey (1996) addresses the same problem. He also presents empirical evidence that recalculations of elementary indices with different index formulae give significant changes in aggregate CPIs, annually by more than two percentage points, in Finland, Sweden, Canada and France. The use of unit values (cf. Subsection 2.3.1.1 for a formal discussion) at the lowest level in a price index is analysed by Balk (1998), which is commonly taken for granted to be an appropriate method of aggregation for prices of homogeneous goods. He tries to answer the questions of the conditions under which a group of goods is sufficiently homogeneous to warrant the use of unit values, and if one needs to restrict the use of unit values to homogenous goods alone. In the context of foreign trade, Silver (2009) criticises the use of aggregate indices which are calculated from unit values at the elementary level. He advocates pure price indices and reveals substantial biases of customs-based unit values: they depend on the structure of quantities and hence, cannot be considered surrogates for survey-based prices.

Szulc (1989) describes the fact that biases at the elementary level are more severe than the pros and cons of the formula at the aggregate level. He finds that if one ignores the particularities of the aggregate index when calculating elementary indices, this might result in surprisingly low differences between different aggregate indices. This is because the indices at the elementary level might not be paying attention to the characteristics of the index formula at the aggregate level, in particular if the same elementary indices are used as building blocks of the aggregate index – no matter which aggregate index should be used. In his 1994 paper he presents numerical evidence for the Canadian CPI that the choice of the elementary index matters the most, particularly in the short term. Dalén (1992, 1995) discusses the impact of the choice of the wrong index formula at the elementary level in the Swedish CPI. Statistics Sweden switched over to the Carli index in January 1990. As soon as April it was replaced by a variant of the geometric index due to the well-known severe upward bias of the Carli index – of more than half a percent in these three months. Using Swedish and Finnish data, he shows in his 1998 paper that the Carli index consistently gives results which are year-on-year two index points and more larger than the Dutot and Jevons indices, while the latter two indices are fairly close to each other. Fenwick (1999) presents evidence that the UK HICP, which is based on the Jevons index at the elementary level, is annually about half a percentage point lower than the national equivalent, the Retail Prices Index, which uses a combination of the Dutot and Carli indices, only because of the different formulae. His main argument for this notable difference is the relative broad item description, leading to aggregation of highly heterogeneous items. Silver and Heravi (2007) show that the difference between the Jevons and Dutot indices is due to different variances in the observed prices at different points in time alone, i.e. these indices will differ if prices exhibit dispersion. From a hedonic regression they derive a heterogeneity-controlled Dutot index and successfully test their approach empirically with scanner data.

2.2 Aggregate Indices

2.2.1 First Principles

At the aggregate level, the target of measurement determines the index concept to be used. This is either the cost of goods (COGI) or the cost of living (COLI). In general, the former case leads to Laspeyres (1871) and Paasche (1874) price indices, while the latter results inter alia in the Fisher (1922) price index – other formulae include the Walsh (1901, 1921) and Törnqvist (1936) price indices.

The Laspeyres price index is the arithmetic mean of price relatives with base period expenditure share weights. Here, p_{ib} and q_{ib} denote the price and quantity, respectively, of the i^{th} good at time $b \in \{0, t\}$.

$$P^L = \sum_{i=1}^{n} \frac{p_{it}}{p_{i0}} \frac{p_{i0}q_{i0}}{\sum_{i=1}^{n} p_{i0}q_{i0}} = \frac{\sum_{i=1}^{n} p_{it}q_{i0}}{\sum_{i=1}^{n} p_{i0}q_{i0}} \tag{2.1}$$

This is the only price index which ensures the principle of pure price comparison (cf. von der Lippe, 2001) over multiple periods by using a fixed basket of goods and which is consistent in aggregation (Subsection 2.2.2 provides a discussion of this property).

For volume measurement, one would opt for the Laspeyres quantity index Q^L, with $Q^L = V/P^P$, where V is the ratio of expenditures at times t and 0 or the value index and P^P is the Paasche price index. One might call Q^L a (volume) index in constant prices (COPI). The Paasche price index is the harmonic mean of price relatives with current period expenditure share weights.

$$P^P = \left(\sum_{i=1}^{n} \left(\frac{p_{it}}{p_{i0}} \right)^{-1} \frac{p_{it}q_{it}}{\sum_{i=1}^{n} p_{it}q_{it}} \right)^{-1} = \frac{\sum_{i=1}^{n} p_{it}q_{it}}{\sum_{i=1}^{n} p_{i0}q_{it}} \tag{2.2}$$

This is the only price index leading to volume measures in constant prices which are consistent in aggregation and purely comparable over multiple periods (Laspeyres principle for the quantity index).

The Fisher price index, among others, is a superlative index. It is defined as the geometric mean of the Laspeyres and Paasche price indices.

$$P^F = \sqrt{\frac{\sum_{i=1}^{n} \frac{p_{it}}{p_{i0}} \frac{p_{i0}q_{i0}}{\sum_{i=1}^{n} p_{i0}q_{i0}}}{\sum_{i=1}^{n} \left(\frac{p_{it}}{p_{i0}} \right)^{-1} \frac{p_{it}q_{it}}{\sum_{i=1}^{n} p_{it}q_{it}}}} = \sqrt{P^L P^P} \tag{2.3}$$

This is the most famous price index approximating the change in the minimum expenditures, which preserve utility at a constant level, owing to changes in (relative) prices (cf. Allen, 1975).

2.2.2 Two-Staged Indices

In what follows, the relation between the elementary and aggregate level of two-staged indices is analysed. Firstly, two-staged indices with the same index formula at both levels are described. Secondly, as in the practice of official statistics, different index formulae are applied at the two levels.

While the Laspeyres and Paasche price indices are consistent in aggregation, the first source of aggregation bias arises from the fact that the Fisher price index is not. This means that the result of a two-staged index calculation does not necessarily coincide with that of a calculation in a single stage. However, as Diewert (1978) shows, superlative indices, such as the Fisher price index, are approximatively consistent in aggregation. Still, the remaining inconsistency can lead to puzzling results. The one-staged index is not necessarily restricted to lie in-between the elementary indices of a two-staged calculation. Even though all elementary indices show decreasing prices, i.e. $P_k^F < 1 \,\forall\, k$ (P_k^F being the Fisher price index for the k^{th} group of goods), the aggregate index can show increasing prices, i.e. $P^F > 1$, and vice versa. Additionally, von der Lippe (2007) proposes the Equality Test and shows that even if all elementary indices are equal, the aggregate index can differ.

Much more severe than this defect of the Fisher price index is the second source of aggregation bias which occurs when statistical offices cannot use a quantity or expenditure-weighted formula at the first stage of the aggregation process. Owing to the unavailability of this information they have to rely on an unweighted index which might not reflect the characteristics of the index formula at the aggregate level. This elementary index bias is equally applicable to the Laspeyres and Paasche price indices as well as to the Fisher price index, no matter which unweighted index is used. A two-staged index with a non-according formula at the elementary level, e.g. $P^{(J)L}$, the Laspeyres price index with Jevons indices as building blocks, can lead to a different conclusion than the true price index. Similarly, as before, one can have decreasing prices with the two-staged index, i.e. $P^{(J)L} < 1$, while the true price index shows increasing prices, i.e. $P^L > 1$, and vice versa. This becomes even worse for the Fisher price index which, in addition, if it is calculated in two stages, can lie outside the bounds of the true Laspeyres and Paasche price indices, i.e. $P^{(J)F} > P^L$ or $P^{(J)F} < P^P$. Both scenarios are due to the fact that the elementary indices may not even be close to the desired target index. Hence, more attention should be paid to the calculation of elementary indices.

2.2.3 Index Theory

From an index theoretical standpoint, there exist four approaches which offer guidance in the choice of an index formula at both the aggregate and elementary levels. These are the economic, axiomatic, stochastic and sampling approaches and they are described below.

The economic approach gives a microeconomic interpretation to consumer's optimising behaviour. Konüs (1924) develops this approach and derives the cost of living index as the solution to a cost minimisation problem. Moreover, he shows that the upper and lower bounds are, in general, the Laspeyres and Paasche price indices: $P^P \leq P^{COLI} \leq P^L$. A COLI measures the change in the minimum expenditures in order to maintain a given level of utility and hence, substitution between goods is permitted. In practice, these indices are approximated by superlative indices, such as the Fisher price index, as discussed in Diewert (1976). This approach assumes that timely information on quantities or expenditures is available.

The Fisher price index is typically the preferred formula from the viewpoint of the axiomatic approach, too. The axiomatic approach states properties which an index formula should desirably fulfil and checks which axioms are actually fulfilled. Eichhorn (1978), and Diewert (1995) discuss this approach to elementary indices in detail. However, the elementary and aggregate levels are treated individually. In order to fill this gap, an integrated approach for two-staged indices would be desirable. The importance of axioms in general depends heavily on the target of measurement (the possible target indices are introduced in Subsection 2.2.1) and, to some extent, on personal preferences.

Selvanathan and Prasada Rao (1994) describe a stochastic approach to index numbers in general. In this approach, the price index is the least squares estimator of a weighted regression of price relatives, enabling the calculation of standard errors and confidence intervals. The idea behind this is that price relatives deviate only randomly from the overall price index. The shortcoming of this approach is that it does not distinguish the fit of the model from the sampling error. The variance of an estimator is rather the expression of the heterogeneity of the price representatives forming the group of goods. One should not take the lowest variance as a measure for determining the most suitable index. Thus, this approach is not designed for judging the adequacy of an index formula. In any case, its main purpose lies in international comparisons.

The sampling approach for elementary indices is presented by Balk (2005, 2008). This approach studies elementary indices as sample estimators of unknown population price indices and the required sampling design for unbiasedness. Under an appropriate sampling scheme, both the Dutot and Carli indices can be justified as sample counterparts of the Laspeyres price index. The appropriate sampling scheme in both cases is "probability proportional to size" (PPS) sampling. For the Dutot index to equal the Laspeyres price index, the price representatives should be sampled according to their quantities in the base period. Should the Carli index equal the Laspeyres price index, the appropriate PPS weights are base period expenditures. This approach has the merit of allowing the achievement of numerical equivalence of an elementary index with the desired target index, i.e. $\mathrm{E}(P^D) = P^L$ or $\mathrm{E}(P^C) = P^L$. The demerit is that the informational requirements (quantities or expenditures) are generally not met by statistical offices. If they were met, the desired target index could be calculated directly.

Thus, owing to the aforementioned limitations, none of these four approaches is followed here but a new, fifth approach is proposed. Although a different path is trodden, the goal which is to be achieved is the same as that of the sampling approach: numerical equivalence. The following statistical approach using power means does not rely on PPS but on simple random sampling (SRS), which requires much less additional information and is easier to implement.

2.3 Corresponding Elementary Indices

2.3.1 Theoretical Foundations

In order to achieve numerical equivalence between an elementary index and an arbitrary aggregate index, a statistical approach is developed. In Subsection 2.3.1.1 it is firstly demonstrated that every weighted index can be expressed one-to-one and onto as a "power mean", as long as the former satisfies the strict mean value property. The power mean represents a whole class of unweighted elementary indices, such as the Carli and Jevons indices. However, an analytical derivation of the concrete power mean of a weighted index, aggregate or elementary, is not possible without further assumptions. Hence, secondly the log-normal distribution is introduced in Subsection 2.3.1.2 and the power means – which correspond to the Laspeyres, Paasche and Fisher price indices, as well as to the Dutot and unit value indices – are related to the distribution's parameters. Although, at that stage, one would be able to numerically calculate elementary indices, corresponding to the desired aggregate ones, the present chapter goes one step further and gives an economic interpretation to the parameters through a partial adjustment model in Subsection 2.3.1.3. Thirdly, the log-normal distribution parameters are related to the price elasticity. Finally, it is shown in the succeeding Subsections 2.3.2 and 2.3.3 that the choice of the elementary indices which correspond to the desired aggregate ones can be based on the price elasticity alone.

2.3.1.1 Power Mean

Right at the very beginning, Lemma 2.1 is needed for the discussion of the problem at the elementary level. Proof for this and all following lemmata and theorems are to be found in Appendix A.

Lemma 2.1. *The price indices of Laspeyres and Paasche as well as the Fisher price index, Equations (2.1), (2.2) and (2.3), respectively, pass the Mean Value Test of Eichhorn and Voeller (1976):*

$$\min\left(\frac{p_{it}}{p_{i0}}\right) \leq P^* \leq \max\left(\frac{p_{it}}{p_{i0}}\right), \tag{2.4}$$

where P^* stands for any of the three price indices. This test says that the price index should be greater than or equal to the lowest price relative and less than or equal to the highest one, with equality if and only if all price relatives are equal.

Given this, the problem of choosing the elementary index corresponding to the Laspeyres, Paasche or Fisher price indices becomes solvable. To this end, it is shown that every *weighted* aggregate index can be written as an *unweighted* power mean of price relatives.

Definition 2.1. Let p_{it}/p_{i0} denote the price relative of the i^{th} good at time t, where $i = 1, 2, \ldots, n$ and $n \geq 2$. Furthermore, all price relatives are assumed to be positive real numbers, $0 < p_{it}/p_{i0} < \infty \, \forall \, i$. Then, their power mean is defined as

$$P^r = \sqrt[r]{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{p_{it}}{p_{i0}}\right)^r}. \tag{2.5}$$

By choosing the appropriate powers r, the resulting power means equal some of the most important elementary indices (cf. Table 2.1). Figure 2.1 exemplifies the typical shape of the power mean as a function of its argument r. Its analytical properties are stated in Lemma 2.2.

Lemma 2.2. *The power mean is a mapping from the affinely extended real numbers $\mathbb{R} \cup (-\infty, +\infty)$ on the closed interval $[P^{min}, P^{max}]$, or technically speaking*

$$P^r : \mathbb{R} \cup (-\infty, +\infty) \rightarrow [P^{min}, P^{max}]. \tag{2.6}$$

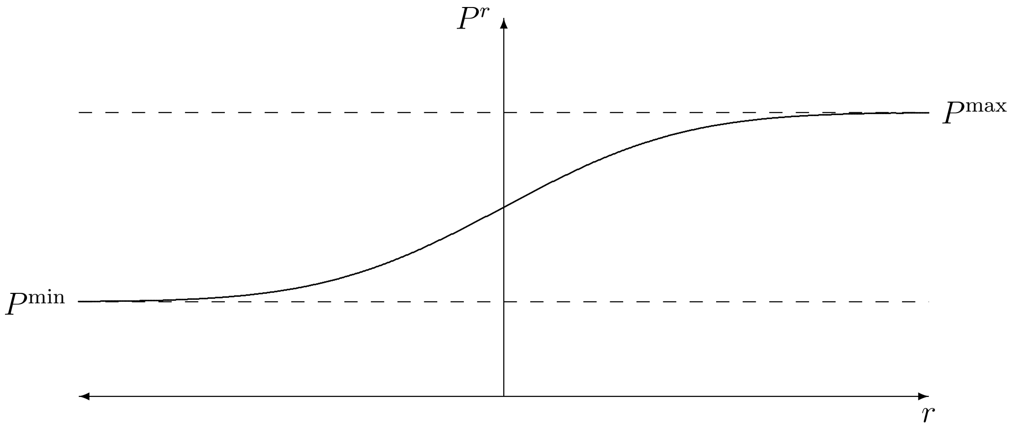


Figure 2.1: Power Mean of Price Relatives

From these intermediate results, the following theorem is deduced.

Theorem 2.1. *If not all price relatives are equal, $\exists\, i \neq j : p_{it}/p_{i0} \neq p_{jt}/p_{j0}$, i.e. the trivial case of perfect homogeneity is neglected, then for any aggregate index P^* that satisfies the mean value property there exists one and only one real r for which the power mean is numerically equivalent,*

$$\exists!\, r \in \mathbb{R} \colon P^r = P^*. \tag{2.7}$$

Theorem 2.1 provides the basis for the following derivation of the corresponding elementary indices in the case of the Laspeyres, Paasche or Fisher price indices as desired aggregate indices. An intuitive interpretation of the theorem goes as follows. The aggregate index P^* lies between the smallest and largest price relative, $P^{\min}$ and $P^{\max}$, respectively. The power mean P^r covers the whole range between these two price relatives. Moreover, it is a continuous function and hence, it has to take on the value of the aggregate index at least once. Uniqueness of the power r is secured through the proposition that not all price relatives are equal and therefore, the power mean is a strictly monotonic increasing function in r.

Table 2.1 depicts some of the most frequently used formulae at the elementary level (cf. Subsection 2.3.3 for the definitions of quadratic means). Appendix C gives a numerical example of how power means equal Laspeyres price indices while Carli indices fail.

Table 2.1: Power Means and Their Formulae

r	Power Mean	Price Index	Formula
-2	reciprocal quadratic	–	$P^r(-2) = \sqrt{n/\sum_{i=1}^n (p_{i0}/p_{it})^2}$
-1	harmonic	Coggeshall (1887)	$P^h = n/\sum_{i=1}^n (p_{i0}/p_{it})$
$0^\dagger$	geometric	Jevons (1863, 1865)	$P^J = \sqrt[n]{\prod_{i=1}^n (p_{it}/p_{i0})}$
1	arithmetic	Carli (1764)	$P^C = \sum_{i=1}^n (p_{it}/p_{i0})/n$
2	quadratic	–	$P^r(2) = \sqrt{\sum_{i=1}^n (p_{it}/p_{i0})^2/n}$

† The Jevons index is the limit of P^r as r approaches zero.

Another very famous formula at the elementary level is the one of Dutot (1738), the ratio of arithmetic mean prices:

$$P^D = \frac{\frac{1}{n}\sum_{i=1}^n p_{it}}{\frac{1}{n}\sum_{i=1}^n p_{i0}}. \tag{2.8}$$

Carruthers et al. (1980) show that this index is related to the Jevons index to the second order via $P^J \approx P^D[1 + \mathrm{Var}(p_0^*)/2 - \mathrm{Var}(p_t^*)/2]$, where $\mathrm{Var}(p_0^*)$ and $\mathrm{Var}(p_t^*)$ are the variances of the relative deviations of the prices from their arithmetic mean in the respective periods: $\nu_{ib} = (p_{ib}/\bar{p}_b) - 1$, $b \in \{0, t\}$. Hence, the two indices will closely approximate each other if the variance of the prices remains constant over time. For this reason, the Dutot index is frequently put on a par with the Jevons index.

Drobisch (1871) proposes another index which is of importance at the elementary level. This is the ratio of unit values or the unit value index:

$$P^{UV} = \frac{\sum_{i=1}^n p_{it}q_{it}/\sum_{i=1}^n q_{it}}{\sum_{i=1}^n p_{i0}q_{i0}/\sum_{i=1}^n q_{i0}}. \tag{2.9}$$

Note that the summation of quantities must be defined and should be economically meaningful. The unit value index is an elementary index in the sense of being a surrogate for a price index (cf. Silver, 2009, and von der Lippe and Mehrhoff, 2009). Using the theorem of von Bortkiewicz (1923), Párniczky (1974) derives criteria under which the unit value index equals the Paasche price index, while Balk (1998) does this for the Fisher price index. They arrive at the following expressions for the ratio of the unit value index to the two indices: $P^{UV}/P^P = 1 + \mathrm{relCov}(p_0, q_t/q_0)$

and $P^{UV}/P^{F} = \sqrt{[1 + \mathrm{relCov}(p_0, q_t/q_0)][1 + \mathrm{relCov}(p_t, q_t/q_0)]}$, respectively, where $\mathrm{relCov}(X, Y) = \mathrm{Cov}(X, Y)/[\mathrm{E}(X)\mathrm{E}(Y)]$. For the unit value index to equal the Paasche price index, at least one of the following criteria has to hold: a) all base period prices have to be equal, b) all quantity relatives have to be equal, or c) there is no correlation between base period prices and quantity relatives. In the case of the Fisher price index, the situations a) and c) have to hold for current period prices as well. For the reason of Lemma 2.3, the unit value index is not a price index in the classical meaning.

Lemma 2.3. *The unit value index in Equation (2.9) violates the mean value property from Equation (2.4).*

However, with respect to its importance in both consumer prices and foreign trade, it will be analysed along with power means and the Dutot index in the next subsection.

2.3.1.2 Log-Normal Distribution

The power r in Subsection 2.3.1.1 cannot be derived analytically without making further assumptions. Based on Theorem 2.2, a closed form solution is provided as to which power corresponds to a given aggregate index as well as to the practically relevant Dutot and unit value indices.

Theorem 2.2. *Under weak assumptions on the underlying data generating process, which are outlined in the proof (cf. Appendix A), prices p_{ib} and quantities q_{ib}, $b \in \{0, t\}$, are jointly log-normally distributed:*

$$\begin{bmatrix} \mathbf{p_i} \\ \mathbf{q_i} \end{bmatrix} \sim \mathcal{LN}\left(\begin{bmatrix} \boldsymbol{\mu}_\mathbf{p} \\ \boldsymbol{\mu}_\mathbf{q} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{p},\mathbf{p}} & \boldsymbol{\Sigma}_{\mathbf{p},\mathbf{q}} \\ \boldsymbol{\Sigma}_{\mathbf{q},\mathbf{p}} & \boldsymbol{\Sigma}_{\mathbf{q},\mathbf{q}} \end{bmatrix} \right). \tag{2.10}$$

Upon this, an explicit formula is derived by which the power can be computed directly from the log-normal distribution parameters. In Subsection 2.3.1.3, these distribution parameters will be linked to the price elasticity.

The characteristic run of the log-normal distribution can be inferred from Figure 2.2.

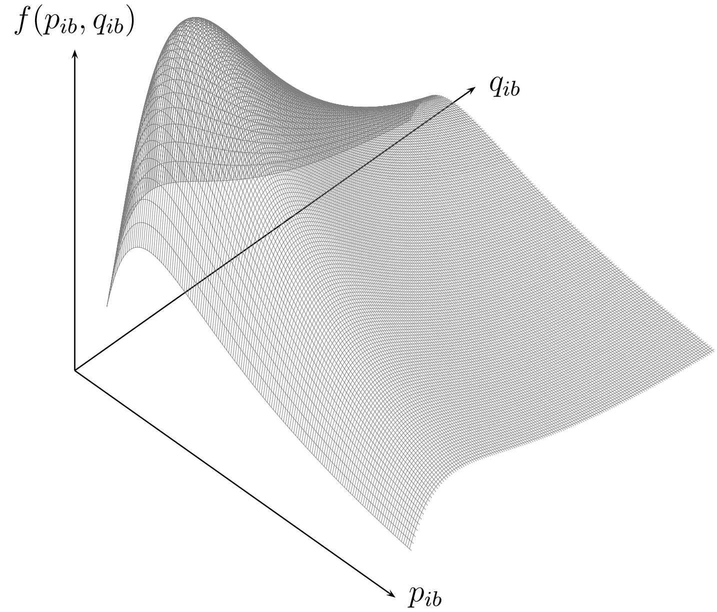


Figure 2.2: Joint Log-Normal Distribution of Prices and Quantities

The assumption of a quadrivariate log-normal distribution of prices and quantities seems reasonable and predecessors are found in the literature. Moulton (1993), and Dalén (1999) use the log-normal distribution assumption for price relatives, while Silver and Heravi (2007) use it for prices in their own right. Note that the latter assumption is a generalisation of the former one. Log-normal distribution of price relatives is a direct consequence of log-normal distribution of prices. In fact, for power means the results are the same from either of the assumptions. However, it would not be possible to analyse the Dutot and unit value indices without the more general assumption.

The link between the power mean, and the Dutot and unit value indices on the one side, and the log-normal distribution parameters on the other side is built in the following theorem.

Theorem 2.3. *The power mean in Equation (2.5) corresponds to the r^{th} root of the r^{th} raw moment of the marginal distribution of price relatives, which is also the log-normal distribution. It follows that*

$$P^r = \exp\left(\mu_{p_t} - \mu_{p_0} + r\frac{\sigma_{p_t}^2 + \sigma_{p_0}^2 - 2\sigma_{p_t,p_0}}{2}\right). \tag{2.11}$$

The Dutot index in Equation (2.8) is the ratio of the first raw moments of the marginal distributions of current and base period prices. One finds that

$$P^D = \exp\left(\mu_{p_t} - \mu_{p_0} + \frac{\sigma_{p_t}^2 - \sigma_{p_0}^2}{2}\right). \tag{2.12}$$

The unit value index in Equation (2.9) is found to be a ratio of ratios. The ratios, either in the current or base period, are those of the first raw product moment of the marginal distribution of prices and quantities and the first raw moment of the marginal distribution of quantities. This results in

$$P^{UV} = \exp\left(\mu_{p_t} - \mu_{p_0} + \frac{\sigma_{p_t}^2 - \sigma_{p_0}^2 + 2\sigma_{p_t,q_t} - 2\sigma_{p_0,q_0}}{2}\right). \tag{2.13}$$

From Theorem 2.3, it can be seen that the Carli index ($r = 1$), unlike the Jevons index ($r \to 0$), is an increasing function of the variance of the price relatives. Hence, a mathematical argument for the upward bias of the Carli index compared with the Jevons index is given through this: the more heterogeneous the goods become at the elementary level, the higher will be the bias.

Theorem 2.4 establishes the link between the Laspeyres and Paasche price indices and the log-normal distribution parameters (cf. Subsection 2.3.3 for the solution in the case of the Fisher price index). Moreover, it firstly gives an exact expression for the power mean corresponding to either of the two price indices, and secondly shows to which power mean the Dutot and unit value indices relate.

Theorem 2.4. *The Laspeyres price index corresponds to the ratio of the first raw product moment of the marginal distribution of current period prices and base period quantities, and the first raw product moment of the marginal distribution of base period prices and quantities. It turns out that*

$$P^L = \exp\left(\mu_{p_t} - \mu_{p_0} + \frac{\sigma_{p_t}^2 - \sigma_{p_0}^2 + 2\sigma_{p_t,q_0} - 2\sigma_{p_0,q_0}}{2}\right). \tag{2.14}$$

The Paasche price index' correspondence is the same as the one of the Laspeyres price index but with the difference that here there are current period quantities instead of base period ones. It becomes

$$P^P = \exp\left(\mu_{p_t} - \mu_{p_0} + \frac{\sigma^2_{p_t} - \sigma^2_{p_0} + 2\sigma_{p_t,q_t} - 2\sigma_{p_0,q_t}}{2}\right). \tag{2.15}$$

Equating P^r from Equation (2.11) with P^L and P^P from Equations (2.14) and (2.15), respectively, and solving for r yields after some algebra:

$$P^r = P^L \iff r_L = \frac{\sigma^2_{p_t} - \sigma^2_{p_0} + 2\sigma_{p_t,q_0} - 2\sigma_{p_0,q_0}}{\sigma^2_{p_t} + \sigma^2_{p_0} - 2\sigma_{p_t,p_0}}, \tag{2.16}$$

$$P^r = P^P \iff r_P = \frac{\sigma^2_{p_t} - \sigma^2_{p_0} + 2\sigma_{p_t,q_t} - 2\sigma_{p_0,q_t}}{\sigma^2_{p_t} + \sigma^2_{p_0} - 2\sigma_{p_t,p_0}}. \tag{2.17}$$

Finally, the Dutot and unit value indices from Equations (2.12) and (2.13), respectively, are related to the power mean as follows:

$$P^r = P^D \iff r_D = \frac{\sigma^2_{p_t} - \sigma^2_{p_0}}{\sigma^2_{p_t} + \sigma^2_{p_0} - 2\sigma_{p_t,p_0}}, \tag{2.18}$$

$$P^r = P^{UV} \iff r_{UV} = \frac{\sigma^2_{p_t} - \sigma^2_{p_0} + 2\sigma_{p_t,q_t} - 2\sigma_{p_0,q_0}}{\sigma^2_{p_t} + \sigma^2_{p_0} - 2\sigma_{p_t,p_0}}. \tag{2.19}$$

2.3.1.3 Partial Adjustment Model

Next, the implied power r of the Laspeyres and Paasche price indices as well as of the Dutot and unit value indices, Equations (2.16) and (2.17) in addition to Equations (2.18) and (2.19), is connected to the price elasticity derived from a partial adjustment model as in Definition 2.2.

Definition 2.2. It is assumed that an equilibrium quantity traded for each good $i = 1, 2, \ldots, n$ and time $b \in \{0, t\}$ exists. This quantity is related to the price of the good, which, in turn, is assumed to be predetermined, and to other, strictly exogenous variables, such as time dummies or a trend. The parameter η_i^q is a panel fixed effect, accounting for unobserved heterogeneity in the data.

$$\ln \overline{q}_{ib} = \alpha + \beta \ln p_{ib} + \mathbf{x_{ib}}\boldsymbol{\delta} + \eta_i^q \tag{2.20}$$

The adjustment to the equilibrium in Equation (2.20) is assumed to be both incomplete and erroneous. This is mirrored by the introduction of lagged quantity and an i.i.d. error term. Here, $\beta^* := (1-\rho)\beta$ denotes the effective price elasticity.

$$\begin{aligned}\ln q_{ib} &= (1-\rho)\ln \overline{q}_{ib} + \rho \ln q_{ib-1} + \varepsilon^q_{ib} \\ &= (1-\rho)\alpha + \beta^* \ln p_{ib} + \rho \ln q_{ib-1} + \mathbf{x_{ib}}(1-\rho)\boldsymbol{\delta} + [(1-\rho)\eta^q_i + \varepsilon^q_{ib}] \end{aligned} \tag{2.21}$$

Prices are assumed to follow a panel AR(1) process:

$$\ln p_{ib} = \gamma_0 + \gamma_1 \ln p_{ib-1} + (\eta^p_i + \varepsilon^p_{ib}). \tag{2.22}$$

Three remarks have to be made regarding the chosen model. First, the implied cross-price elasticity in Equation (2.20) is zero. Second, the underlying equilibrium price elasticity β is attenuated by sluggish adjustment of quantities. Third, owing to the problem of identification with observed data on prices and quantities, the estimated effective price elasticity β^* has to be understood as being the one of the supply-demand equilibrium rather than the one of demand. As the focus of this chapter is on the effective price elasticity only, it is referred to simply as the price elasticity in what follows.

Using Equations (2.21) and (2.22), the covariance matrices can be derived subject to the model parameters. The results are collected in Theorem 2.5.

Theorem 2.5. *The covariance matrices $\mathbf{\Sigma_{p,p}}$ and $\mathbf{\Sigma_{p,q}} = \mathbf{\Sigma'_{q,p}}$ of the log-normal distribution as given in Equation (2.10) are as follows (the elements of $\mathbf{\Sigma_{q,q}}$ do not appear in the calculation of the power r):*

$$\mathbf{\Sigma_{p,p}} = \begin{bmatrix} \sigma^2_{p_t} & \sigma_{p_t,p_0} \\ \sigma_{p_t,p_0} & \sigma^2_{p_0} \end{bmatrix} = \sigma^2_p \begin{bmatrix} 1 & \gamma_1^t \\ \gamma_1^t & 1 \end{bmatrix}, \tag{2.23}$$

$$\mathbf{\Sigma_{p,q}} = \begin{bmatrix} \sigma_{p_t,q_t} & \sigma_{p_t,q_0} \\ \sigma_{p_0,q_t} & \sigma_{p_0,q_0} \end{bmatrix} = \beta^* \sigma^2_p \begin{bmatrix} \frac{1}{1-\rho\gamma_1} & \frac{\gamma_1^t}{1-\rho\gamma_1} \\ \left(\frac{\gamma_1^t - \rho^t}{1-\frac{\rho}{\gamma_1}} + \frac{\rho^t}{1-\rho\gamma_1}\right) & \frac{1}{1-\rho\gamma_1} \end{bmatrix}. \tag{2.24}$$

In the derivation of Equations (2.23) and (2.24) use was made of the weak stationarity assumption, especially of stationarity in covariance.

2.3.2 Laspeyres and Paasche Price Indices

It is shown that the solution to the problem of elementary indices that correspond to a desired aggregate index depends on the empirical correlation between prices and quantities. In particular, the power r is a function of the price elasticity alone. The succeeding theorem summarises the results for the Laspeyres and Paasche price indices (again, cf. Subsection 2.3.3 for the solution in the case of the Fisher price index) as well as the Dutot and unit value indices.

Theorem 2.6. *Combining the equations relating the power mean to the log-normal distribution parameters, Equations (2.16), (2.17), (2.18) and (2.19), with those relating the log-normal distribution parameters to the model coefficients, Equations (2.23) and (2.24), gives the final results:*

$$r_L = -\beta^* \frac{1}{1-\rho\gamma_1} \approx -\beta^*, \tag{2.25}$$

$$r_P \xrightarrow{t\to\infty} \beta^* \frac{1}{1-\rho\gamma_1} \approx \beta^*, \tag{2.26}$$

$$r_D = 0, \tag{2.27}$$

$$r_{UV} = 0. \tag{2.28}$$

From Theorem 2.6, the general results for the power mean are as follows. A power mean with power r equal to minus the price elasticity ($-\beta^*$) yields approximately the same result as the Laspeyres price index. Hence, if the price elasticity is minus one, for example, the power must equal one and the Carli index (cf. Table 2.1) at the elementary level will correspond to the Laspeyres price index as target index. This can be seen in the simplest form from the following example: from $q_{i0} = \bar{q}_0/p_{i0}$, where $\bar{q}_0$ is an arbitrary constant, follows $P^L = [\sum_{i=1}^n p_{it}(\bar{q}_0/p_{i0})]/[\sum_{i=1}^n p_{i0}(\bar{q}_0/p_{i0})] = \sum_{i=1}^n (p_{it}/p_{i0})/n = P^C$. However, if the Paasche price index should be replicated, the power of the power mean must equal the price elasticity, in the above example minus one. Thus, the harmonic index gives the same result and therefore, in this case it should be used at the elementary level.

Under the assumption of stationarity in covariance (cf. Subsection 2.3.1.3), the Dutot and unit value indices both equal the Jevons index. But if price dispersion takes place in reality, violating this assumption, the indices will differ. This is even more the case for the unit value index than for the Dutot index.

The formulae of power means which correspond to the Laspeyres and Paasche price indices resemble the definition of the constant elasticity of substitution (CES) price index laid down in Appendix B. But this similarity is only superficial.

2.3.3 Fisher Price Index

The Fisher price index is derived from the Laspeyres and Paasche price indices as their geometric mean. Owing to the symmetry of the power means which correspond to the Laspeyres and Paasche price indices, a quadratic mean corresponds to the Fisher price index. In Definition 2.3 the properties of quadratic means in general are presented.

Definition 2.3. A quadratic mean of price relatives of order q is defined as follows:

$$P^q = \left(\frac{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{p_{it}}{p_{i0}}\right)^{\frac{q}{2}}}{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{p_{it}}{p_{i0}}\right)^{-\frac{q}{2}}} \right)^{\frac{1}{q}} \tag{2.29}$$

The index defined by Equation (2.29) is symmetric, i.e. $P^q = P^{-q} = P^{|q|}$. Furthermore, it is either increasing or decreasing in $|q|$, depending on the data. Both characteristics can also be seen from Figure 2.3. Note that a quadratic mean of order q, P^q, should not be mistaken for the quadratic index, $P^r(2)$ (cf. Table 2.1).

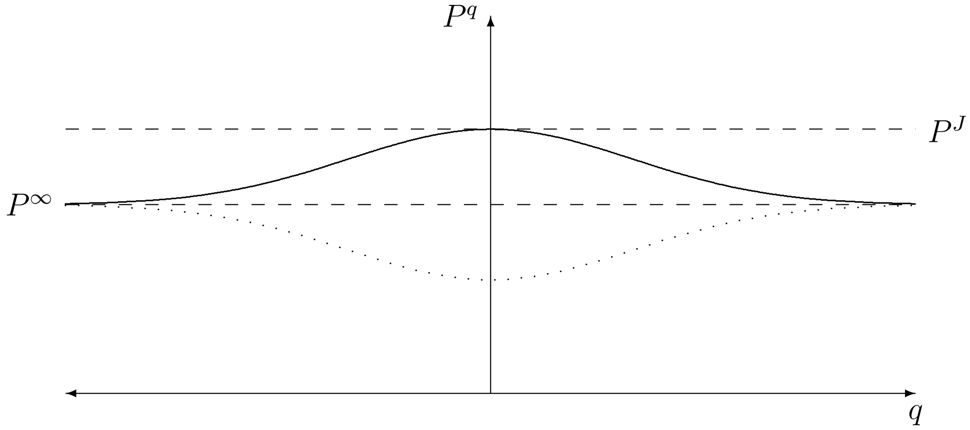


Figure 2.3: Quadratic Mean of Price Relatives

Dalén (1992), and Diewert (1995) show via a Taylor series expansion that all quadratic means approximate each other to the second order. However, as Hill (2006) demonstrates, the limit of P^q if q diverges is $P^\infty = \sqrt{P^{\min} P^{\max}}$. He concludes that quadratic means are not necessarily numerically similar.

For $q \to 0$ the quadratic mean becomes the Jevons index. For $q = 1$ a hybrid index results, which was first described by Balk (2005, 2008) and independently devised by Mehrhoff (2007) as a linear approximation to the Jevons index by crossing the implicit quantities of the Carli and harmonic indices, which explains the name. Implicit quantities are derived by equating the Carli index to the Laspeyres price index and the harmonic index to the Paasche price index; these are the inverses of base and current period prices, respectively (cf. Subsection 2.3.2). Lastly, one arrives at the CSWD index (Carruthers, Sellwood and Ward, 1980, and Dalén, 1992) for $q = 2$, which is the geometric mean of the Carli and harmonic indices. Table 2.2 contrasts these indices.

Table 2.2: Quadratic Means and Their Formulae

q	Quadratic Mean	Formula
$0^\dagger$	Jevons	$P^J = \sqrt[n]{\prod_{i=1}^n (p_{it}/p_{i0})}$
1	Hybrid	$P^H = \sum_{i=1}^n \sqrt{(p_{it}/p_{i0})} / \sum_{i=1}^n \sqrt{(p_{i0}/p_{it})}$
2	CSWD	$P^{CSWD} = \sqrt{\sum_{i=1}^n (p_{it}/p_{i0})} / \sqrt{\sum_{i=1}^n (p_{i0}/p_{it})}$
3	cubic	$P^q(3) = \sqrt[3]{\sum_{i=1}^n \sqrt{(p_{it}/p_{i0})^3}} / \sqrt[3]{\sum_{i=1}^n \sqrt{(p_{i0}/p_{it})^3}}$
4	quartic	$P^q(4) = \sqrt[4]{\sum_{i=1}^n (p_{it}/p_{i0})^2} / \sqrt[4]{\sum_{i=1}^n (p_{i0}/p_{it})^2}$

† The Jevons index is the limit of P^q as q approaches zero.

Applying the preceding definitions gives the final result which is stated in Theorem 2.7.

Theorem 2.7. *A quadratic mean of order two times the absolute price elasticity corresponds to the Fisher price index:*

$$P^F \approx \sqrt{\left(\frac{1}{n}\sum_{i=1}^n \left(\frac{p_{it}}{p_{i0}}\right)^{-\beta^*}\right)^{-\frac{1}{\beta^*}} \left(\frac{1}{n}\sum_{i=1}^n \left(\frac{p_{it}}{p_{i0}}\right)^{\beta^*}\right)^{\frac{1}{\beta^*}}} = P^q(2|\beta^*|). \qquad (2.30)$$

The approximate equality in Equation (2.30) follows from Equations (2.25) and (2.26) in conjunction with Equation (2.5).

2.4 Findings in Foreign Trade Statistics

2.4.1 Data Description

The statistical approach, developed in the preceding section, is applied to real data in this section. In other words, the methodology outlined here is illustrated by an example. At first, an overview of the data and their properties is given. The stationarity assumption, which was used in the derivation of the approach, is justified on empirical grounds. Then, the model is estimated for various goods. The results are presented in terms of the model parameters and as regards the implications for price statistics. Eventually, the model based approach is tested for its robustness in a case study.

An application to scanner data for homogeneous goods would be suited best because information on both prices and quantities at the elementary level is necessary which scanner data would provide. Unfortunately, this kind of data are not available for the German CPI. Hence, as an empirical application, data from German foreign trade statistics are analysed as an alternative. The source of these data is the German Federal Statistical Office. At the time of frontier crossing, movements of goods in special trade are to be reported for statistical purposes; with member states of the European Union in the Intrastat system, and with non-member states via the customs' Single Administrative Document (EC, 2006). Declarations are to be made according to the Commodity Classification for Foreign Trade Statistics and consist inter alia of the goods' values and quantities, the latter generally in terms of the weights. Based on these declarations, albeit not derived from homogeneous goods, unit values are calculated at the elementary level as $\tilde{p}_{ib} = (\sum_{i=1}^{n} p_{ib}q_{ib})/(\sum_{i=1}^{n} q_{ib})$, $b \in \{0, t\}$, which, in turn, form the basis for the succeeding analysis.

Owing to the nature of collected data, their structure is repeated cross-sections rather than a panel. Repeated cross-sections arise by independent cross-sectional surveys at consecutive points in time. Unlike in price statistics, it is not ensured in foreign trade statistics that the same goods are observed over time. The coverage of the universe of goods is time-varying and it is not possible to establish a one-to-one correspondence between goods over time. In this case, Deaton (1985) suggests estimation to be performed on a pseudo panel. This is averaging the data within a cohort, where a cohort is a group of goods sharing common characteristics and every good belongs to one group and one group only which is the same over time. Here, unique transactions are aggregated at the lowest level available, that is their reporting level: the eight-digit code of the Commodity Classification. These lower level aggregates are the individual observations which are nested at the four-digit code level to form an upper level aggregate.

The data set covers 1,264 pseudo panels (nests) consisting of 12,948 groups of goods (cohorts), for exports as well as for imports, and a total of 1,839,384 observations over the period January 2000 to December 2007. Only goods measured in kilograms – these are about three-quarters of all goods – are included in the analysis. The data, unit values in €1,000 per 100 kg (hereafter "prices") and weights in 100 kg (hereafter "quantities"), are transformed into their natural logarithms. Although the goods at the elementary level are not homogeneous, they are treated as if they were for the following analysis.

2.4.2 Regression Results

As weak stationarity of prices and quantities is assumed in the derivation of corresponding elementary indices (cf. Subsection 2.3.1.3), panel unit root tests are performed prior to estimation in order to test the validity of this assumption. In particular, these are the test of Levin, Lin and Chu (LLC, 2002), Breitung (2000), Im, Pesaran and Shin (IPS, 2003), augmented Dickey and Fuller (ADF, 1979), and Phillips and Perron (PP, 1988). The first two assume a common unit root process under the null hypothesis and no unit root under the alternative. The last three, by contrast, test the null hypothesis of an individual unit root process against the alternative of some cross-sections without a unit root. The latter two tests for panel data are derived as a combination of their time series variants using the results of Fisher (1925). Included in the test specification are individual effects and individual linear trends. Lag lengths, if necessary, are selected automatically based on the Schwarz information criterion; if applicable, the spectral estimator's bandwidth is selected according to Newey and West (1994) using the Bartlett kernel.

As can be seen from Table 2.3, the tests show stationarity of both prices and quantities for almost all panels in exports as well as in imports. Throughout, quantities perform better than prices, and exports and imports do equally well. That not all of them are stationary is largely due to non-unity power of the tests. Thus, the issue of (co-)integration can safely be ignored for the remainder of the analysis.

Table 2.3: Percentages of Stationary Panels at the 5% Significance Level

	Exports		Imports	
Test	Prices	Quantities	Prices	Quantities
LLC	89.81%	94.82%	90.46%	93.07%
Breitung	84.47%	90.88%	85.66%	92.26%
IPS	93.34%	97.70%	92.50%	97.72%
ADF	93.51%	97.78%	93.07%	98.04%
PP	96.87%	98.72%	96.88%	99.36%

The price elasticity β^* is estimated in the framework of the log-linear partial adjustment model given in Equation (2.21) by means of dynamic panel data one-step system GMM (Arellano and Bover, 1995, and Blundell and Bond, 1998). Neither time dummies nor a deterministic trend are included. Prices are assumed to be predetermined and are instrumented accordingly. The instrument set is collapsed in order to reduce the instrument count.

The overall results are fairly robust to different specifications of the model (inclusion of dummies or a trend), choice of instruments (limited lag depth) and estimation methods (fixed effects or difference GMM). Thus, only results which are derived from the above set-up are reported.

After adjusting for outliers, 1,246 panels in exports and 1,249 in imports remain. The distribution of the price elasticity in exports and imports can be gathered from Figure 2.4. The histograms show positive excess kurtosis, or leptokurtosis, for exports as well as for imports. Compared with the associated normal distribution, the peak around the mean is more pronounced, i.e. there is a higher probability of values near the mean, and the tails are fatter, i.e. there is a higher probability of extreme values. However, the distributions look both quite unimodal and symmetric. The distribution of imports lies slightly more to the right than the one of exports.

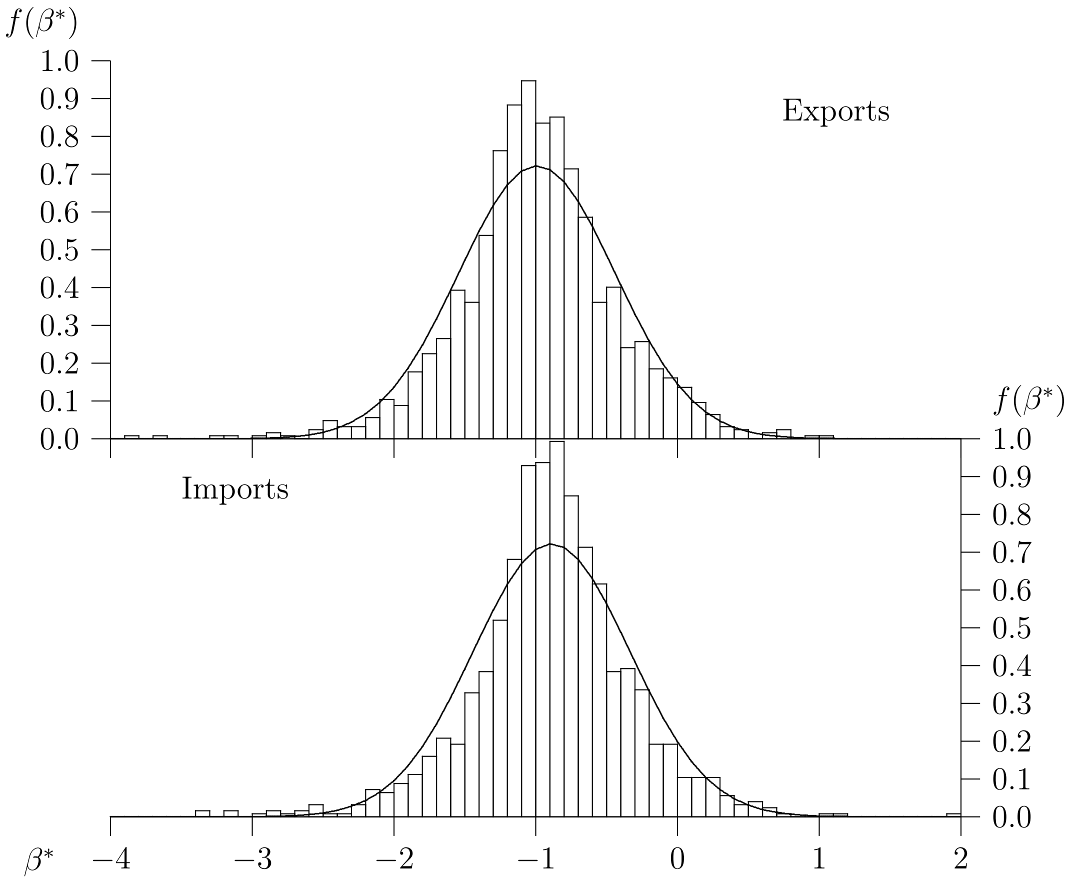


Figure 2.4: Density Histogram (Bin Width = 0.1) and Normal Density Plot of β^*

The most important descriptive summary statistics are collected in Table 2.4. The average price elasticity (β^*) for exports is −0.99, ranging from −3.9 to 1.1. Adjustment to the equilibrium is strong with the adjustment parameter $(1 - \rho)$ being 0.80 on average, lying in the range from 0.0 (no adjustment) to 1.6 (over-adjustment). The goodness-of-fit measure (Pseudo-R^2) is high at 0.51 on average, covering the whole range from 0.0 to 1.0. The results for imports are almost the same with the notable difference of the average price elasticity, which is −0.89, i.e. a significant 0.1 point lower than for exports. Adjustment and goodness-of-fit are as strong as for exports. Yet results for imports are less stable than those for exports. This is due to higher heterogeneity of observed data owing to the large number of different countries from which German companies import goods.

Table 2.4: Descriptive Summary Statistics of the Partial Adjustment Model

	Exports			Imports		
Statistic	β^*	$1-\rho$	Pseudo-R^2	β^*	$1-\rho$	Pseudo-R^2
Mean	−0.9911	0.8014	0.5060	−0.8877	0.8071	0.5116
Variance	0.3055	0.0466	0.0775	0.3055	0.0425	0.0748
Minimum	−3.8923	−0.0157	0.0001	−3.3727	0.0491	0.0000
Maximum	1.0826	1.6344	0.9961	1.9547	1.4127	0.9850

As a goodness-of-fit measure, a Pseudo-$R^2 = \mathrm{Corr}^2(\ln q_{ib}, \ln \hat{q}_{ib})$ is used. This is the squared coefficient of correlation between observed and fitted values with the obvious interpretation of explained variance of a regression of observed on fitted values and an intercept.

Persistence of the process of prices given in Equation (2.22) is relatively low; on average, the autoregressive parameter γ_1 is 0.17 for exports and 0.19 for imports, thus rendering the simplification of Theorem 2.6 valid.

More important than the regression results themselves are their implications for price statistics. These are summarised in Tables 2.5 and 2.6, respectively, in terms of the proportions to which each of the elementary indices corresponds to the Laspeyres, Paasche and Fisher price indices for panels with at least two groups of goods. While the first set of figures relates to the number of panels, the second set mirrors their monetary value. The classification of the panels to elementary indices is based on rounding the powers r and q, respectively, to the closest integer. This procedure is illustrated in Figure 2.6.

For the Laspeyres price index as the desired aggregate index, 70% of the panels in exports and 72% in imports imply the use of the Carli index at the elementary level. This means that if one wants to calculate a Laspeyres price index at the aggregate level, the Carli index will yield approximately the same result at the elementary level in these panels (as it is shown in an example in Subsection 2.3.2). Regarding trade values, these figures reduce to 62% and 66%, respectively. The Jevons index performs best at the first stage in 14% of the panels in exports and 17% in imports with much higher shares with respect to trade value, i.e. 29% and 28%, respectively. In 15% of the panels in exports and 10% in imports, the quadratic index is desirable at the lower level of aggregation; trade value shares here are 7% and 5%, respectively. Shares missing to 100% reflect other indices.

Table 2.5: Elementary Indices Corresponding to a Laspeyres Price Index

		Panels		Trade Values	
r	Price Index	Exports	Imports	Exports	Imports
0	Jevons	14%	17%	29%	28%
1	Carli	70%	72%	62%	66%
2	quadratic	15%	10%	7%	5%

If the Paasche price index is taken as the desired aggregate index, the corresponding power means are inverted: instead of the Carli index, the harmonic index, and, instead of the quadratic index, the reciprocal quadratic index have to be used. As mentioned before, if the Jevons index corresponds to the Laspeyres price index, it does so for the Paasche price index, too.

If the desired aggregate index is chosen to be the Fisher price index, the results are as follows. The use of the Jevons index is suggested by 6% of the panels in exports and 7% in imports; with respect to trade values these shares increase to 20% and 17%, respectively. The hybrid index is found to be superior in 21% of the panels in exports (trade values: 19%) and 28% (25%) in imports. 46% of the panels in exports (48%) and 44% in imports (43%) favour the CSWD index. A quadratic mean of cubic order should be used for 21% of the panels in exports (9%) and 15% in imports (12%). For a quadratic mean of quartic order the figures are 6% (3%) and 4% (2%), respectively. Again, quintic and higher orders make up shares missing to 100%.

Table 2.6: Elementary Indices Corresponding to a Fisher Price Index

		Panels		Trade Values	
q	Price Index	Exports	Imports	Exports	Imports
0	Jevons	6%	7%	20%	17%
1	Hybrid	21%	28%	19%	25%
2	CSWD	46%	44%	48%	43%
3	cubic	21%	15%	9%	12%
4	quartic	6%	4%	3%	2%

All in all, different elementary indices should be applied to each panel in order to approach the Laspeyres, Paasche or Fisher price index as closely as possible.

2.4.3 A Case Study

In Subsection 2.4.2 neither the Dutot nor the unit value index could be analysed. The presumption that these indices will differ from the Jevons index due to price dispersion (as discussed in Subsection 2.3.2) can only be tested with sufficient data. The intention of the following case study is firstly to discuss the empirical behaviour of these two indices, and secondly to test the results from Subsection 2.4.2 for their robustness using the empirical methodology of Shapiro and Wilcox (1997). Exports of passenger cars are chosen as an example. With an export value of more than €100 billion in 2007, the more than five million cars exported make up more than 10% of trade value of all exported goods. The panel 8703 (four-digit code of the Commodity Classification: motor cars and other motor vehicles principally designed for the transport of persons, including station-wagons and racing cars) consists of 21 groups of goods and 1,895 observations of trade values and quantities in terms of both weight and number (the average 2007 car weights about 1.5 tonnes). The data set ends in 2007 and hence, is not affected by the recent financial crisis which has hit car makers hard around the world.

Both prices and quantities with respect to weight as well as number pass all of the five panel unit root tests of Subsection 2.4.2 at any conventional significance level. The partial adjustment model is robust to the specification of quantities as either weight or number as the results in Table 2.7 indicate. Irrespective of the definition of quantities, the results are virtually the same. The price elasticity for exports is close to zero and insignificant. The adjustment to the equilibrium, at 81%, is as strong as the Pseudo-R^2 is high at 95%. Hence, the Jevons index, corresponding to a price elasticity of zero, seems appropriate for all three, the Laspeyres, Paasche and Fisher, price indices.

Table 2.7: Partial Adjustment Model for Passenger Cars

	Weight			Number		
Statistic	β^*	$1-\rho$	Pseudo-R^2	β^*	$1-\rho$	Pseudo-R^2
Parameter	0.0469	0.8078	0.9572	-0.0277	0.8111	0.9525
Standard Error	0.0303	0.0317	—	0.0290	0.0317	—

Explanations for the counterintuitive result of equal car sales irrespective of prices are threefold, two of which are technical and one is economic. First, the price elasticity derived from the partial adjustment model is the effective one, i.e. lagged adjustment to the equilibrium lowers the absolute value of the price elasticity. Second, the estimated price elasticity should not be mistaken for the one of demand owing to the problem of identification (as explained in Subsection 2.3.1.3). Third, car makers might be wanting to hold sales stable by compensating for exchange rate fluctuations, and might thus be willing to accept short-term reductions in their profits.

After balancing the panel, 15 groups of goods remain for robustness testing of the regression results. Given the strongly balanced feature of this new panel, time series of the desired aggregate indices can be directly calculated. The power mean P^r, Equation (2.5), which minimises the root mean squared error (RMSE) to the desired aggregate index P^*, that is either the Laspeyres or Paasche price index, Equations (2.1) and (2.2), respectively, is found by non-linear least squares:

$$r_{\min} = \arg\min_{r\in\mathbb{R}} \sqrt{\frac{1}{t}\sum_{b=1}^{t}(P_b^* - P_b^r)^2}. \tag{2.31}$$

Analogously to Equation (2.31), one finds the quadratic mean P^q, Equation (2.29), which minimises the root mean squared error to the Fisher price index, Equation (2.3).

In Table 2.8 the outcomes of the partial adjustment model and non-linear optimisation are compared, along with the corresponding power means of the Dutot and unit value indices. The findings do not change qualitatively. In fact, the deviation from the symmetry proposition is insignificant and it turns out that the regression results coincide with the direct calculation of the power mean. The use of the Jevons index is justified. Note that the linear regression is based on a panel data set of 1,817 observations, while the non-linear direct calculation is based on a time series of 95 observations, which makes the latter more prone to erratic behaviour.

Table 2.8: Partial Adjustment Model Compared to Non-Linear Optimisation

Statistic	β^*	r_L	r_P	q	r_D	r_{UV}
Expectation	—	$-\beta^*$	β^*	$2\lvert\beta^*\rvert$	0	0
Parameter	0.0469	0.3000	−0.4224	0.0000	0.2508	0.9920
Standard Error	0.0303	0.1399	0.1453	†	0.0428	0.1528
Pseudo-R^2	0.9572	0.1957	0.1437	0.1548	0.9328	0.2754
RMSE	—	0.0404	0.0388	0.0389	—	—

† Standard error of q is not stable with respect to different initial values.

That the implied power means of the Dutot and unit value indices are significantly off their expectations can be explained with recourse to Equations (2.18) and (2.19). While scanner data in a CPI may be well-behaved in terms of their covariance stability, things are different in an export or import price index (cf. Subsections 2.1.2 and 2.3.1.1 for empirical and theoretical evidence, respectively). Unlike scanner data, the basis of index calculation is not a panel but rather repeated cross-sections with time-varying coverage of the universe of goods, i.e. new goods are introduced while others disappear. Thus, the relative broad item description in foreign trade is likely to cause heterogeneity to increase over time (cf. Subsection 2.3.2 for a discussion of this issue). Neither the variance of prices nor the concurrent covariance between prices and quantities is stable over time. While the variance is increasing, the covariance is decreasing, which explains the gap between the indices and to their expectations.

Table 2.9: RMSEs of Elementary Indices to Desired Aggregate Ones

Estimator	Laspeyres (r_L)	Paasche (r_P)	Fisher (q)
β^*	0.0418	0.0413	0.0389
$r_{\min}$ / $q_{\min}$	0.0404	0.0388	0.0389
Jevons Index	0.0415	0.0408	0.0389
Dutot Index	0.0338	0.0387	0.0338
Unit Value Index	0.0418	0.0650	0.0531

However, when allowing for non-power means as the Dutot and unit value indices, the findings change slightly. For all three target indices, the Dutot index has a lower RMSE than the Jevons index, which shows a lower RMSE than the unit value index. This is depicted in Table 2.9, which compares the RMSEs of the respective elementary indices as estimators of the desired aggregate ones. Nonetheless, the Dutot and Jevons indices are numerically very close.

In Figure 2.5 the time series of the Dutot, Jevons and unit value indices as estimators of the Fisher price index are drawn on the semi-logarithmic scale with base month January 2000 = 100. Both the Dutot and Jevons indices are similar to the Fisher price index and to each other. This was to be expected from the regression results as well as non-linear optimisation (cf. Table 2.8). Owing to missing expenditure share weights, these approximations to the Fisher price index are the closest that one can get. The unit value index is much more volatile (cf. Table 2.9) and lies well above the Fisher price index, although it is fairly close at the beginning of the time series. This was to be expected as well given the aforementioned time-varying variances of prices and covariances between concurrent prices and quantities.

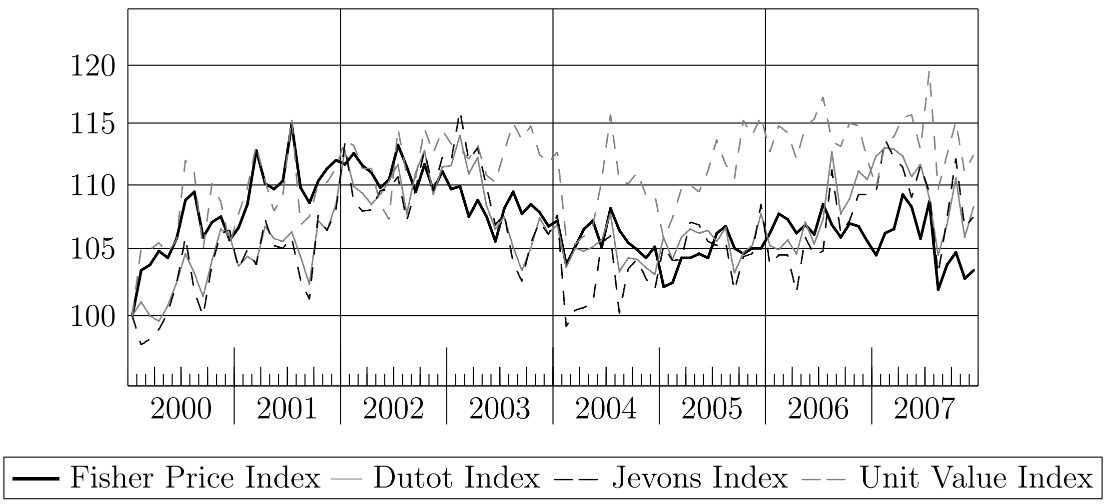


Figure 2.5: Elementary Indices as Estimators of the Fisher Price Index

2.5 Conclusion

2.5.1 Summary

This chapter addresses the problem of index calculation at the elementary level, where no expenditure share weights are available. The question of "which index formula at the elementary level corresponds to the characteristics of the index at the aggregate level?" is dealt with. A statistical approach is proposed which theoretically allows the achievement of numerical equivalence of an elementary index with the desired aggregate index – in this instance, the Laspeyres, Paasche or Fisher price index. Based on "power means" and the assumption of joint log-normal distribution of prices and quantities, it is shown that the solution depends on the price elasticity alone, which is derived from a partial adjustment model. Thus, a feasible framework is provided which aids the choice of the corresponding elementary index. The results are graphically produced in Figure 2.6. If, for example, the price elasticity β^* is minus one, the Carli index corresponds to the Laspeyres price index, the harmonic index to the Paasche price index and the CSWD index to the Fisher price index.

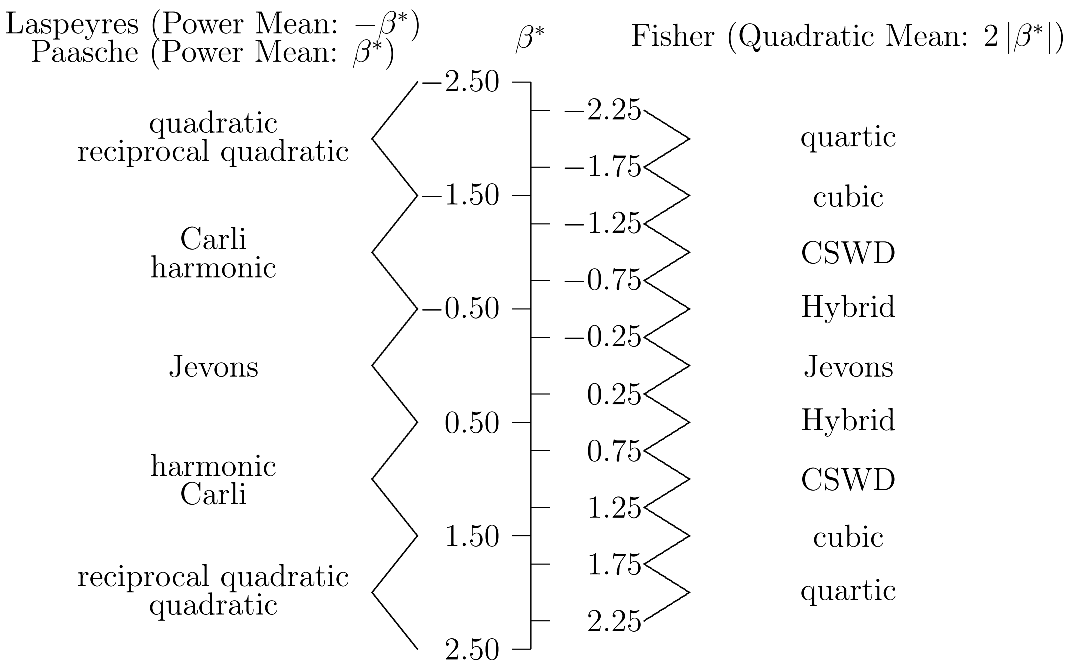


Figure 2.6: Overview of Corresponding Elementary Indices

From an empirical application to German foreign trade statistics, it can be seen that the choice of the elementary index does matter (cf. Figure 2.5). The choice itself depends on the characteristics of prices and quantities. Therefore, depending on the price elasticity, different elementary indices should be applied to each group of goods in order to approach the Laspeyres, Paasche or Fisher price index as closely as possible. While not relying on axiomatic considerations, this chapter finds notable empirical differences between different elementary indices and aggregate indices formed from them. Furthermore, the results indicate that a range of elementary indices should be applied in the calculation of price indices. This is in line with the findings of other authors (cf. the review of the empirical literature in Subsection 2.1.2). In particular, the Carli index performs remarkably well at the elementary level of a Laspeyres price index, corresponding to a price elasticity of minus one. Sometimes, it is argued that the Carli index is upward biased. However, this holds only in comparison with the Jevons index. Yet, the comparison in question is not with another elementary index but with the desired aggregate index. So, it may be the case that the Carli index is unbiased or even downward biased compared with the Laspeyres price index (cf. Subsection 2.3.1.2 for the discussion of the Carli index' upward bias).

2.5.2 Outlook

Two possible applications of the approach outlined in this chapter arise immediately after a decision has been taken on which aggregate index is desired. Firstly, index calculation can be rendered more precise if different elementary indices are applied to each group of goods, reflecting their specific price elasticities. At least for prominent groups of goods with high expenditure shares, studies on the price elasticity should be available. This will drive down biases of official price indices. In fact, the desired aggregate index can be approximated by using appropriate elementary indices. Secondly, for different purposes – either price or volume measurement – different elementary indices should be calculated for the same data. This means that if the Carli index is applied as the single formula at the elementary level of a Laspeyres price index, implying a price elasticity of minus one, the harmonic index must be used at the elementary level of a Paasche price index.

Still, this is in contrast to the current practice as regards foreign trade in Germany, where the Carli index is used at the elementary level in both price statistics and volume measurement in national accounts. The former task is achieved via the Laspeyres price index, while the latter results in an implicit deflator in the form of the Paasche price index.

An application of this approach to scanner data in a CPI would be worthwhile. Scanner data in its most familiar form are collected at the checkouts of retail stores by the scanning of bar codes. Thus, they provide a census of all transactions rather than a sample. Furthermore, they are collected continuously and provide simultaneous information on both prices and quantities, unlike discrete surveys of prices alone. Lastly, qualitative information may be linked to scanner data, allowing for hedonic adjustment. The foreign trade application of this chapter and the prospective study of scanner data are different subject matters. In foreign trade statistics, the data are intermediately aggregated and unit values are used, which are neither seasonally nor quality adjusted, rather than observed purchase prices. Disaggregate scanner data allow the calculation of unbiased price indices and hence, a more thorough analysis based on them might change the results.

Appendix A: Proof of Lemmata and Theorems

Proof of Lemma 2.1. The Laspeyres and Paasche price indices are basket indices, i.e. they are a weighted means of price relatives, either arithmetic with base period expenditure share weights or harmonic with current period expenditure share weights. Either way, the weights $\omega_i > 0$ sum up to one, $\sum_{i=1}^{n} \omega_i = 1$ and the proposition follows. This holds as well for the Fisher price index as it is the geometric mean of the Laspeyres and Paasche price indices. □

Proof of Lemma 2.2. That the limits towards $r \to \pm\infty$ are the maximum and minimum, respectively, can be shown by straightforward algebraic manipulations. The geometric index as the limit towards $r \to 0$ is found via a Taylor series expansion. From this the proposition follows. □

Proof of Theorem 2.1. From Lemma 2.1, it follows that the aggregate index P^* lies between the smallest and largest price relative, $P^{\min}$ and $P^{\max}$, respectively. To reiterate, the exclusion of the trivial case of perfect homogeneity ensures that the mean value property is fulfilled in its strict from.

$$P^{\min} < P^* < P^{\max}$$

That the power mean P^r is continuous on its whole domain follows from Lemma 2.2. Moreover, it covers the whole range between the smallest and largest price relative as its co-domain. Over and above that, the strict mean value property leads to r being real, $r \in \mathbb{R}$.

$$P^{\min} < P^r < P^{\max}$$

In addition, from the intermediate value theorem (under a continuous function the image of a connected space is connected), it follows that the power mean takes on all values of its co-domain, of which the aggregate index is an element, at least once. Hence, the image equals the co-domain.

$$P^* \in (P^{\min}, P^{\max}) \leftarrow \mathbb{R} : P^r$$

Eventually, the uniqueness of the power r is secured through the proposition that not all price relatives are equal, and with Jensen's inequality it can be shown that the power mean is a strictly monotonic increasing function.

$$P^s > P^r \,\forall\, s > r$$

From this, it follows that the power mean is bijective and therefore an inverse function exists. $\square$

Proof of Lemma 2.3. If one writes the unit value index in its price relatives form, $\sum_{i=1}^{n}(p_{it}/p_{i0})\omega_i$, the assigned weights $\omega_i = (p_{i0}q_{it}/\sum_{i=1}^{n} q_{it})/(\sum_{i=1}^{n} p_{i0}q_{i0}/\sum_{i=1}^{n} q_{i0})$ do not necessarily sum up to unity. This contradiction proves the proposition. $\square$

Proof of Theorem 2.2. The processes of prices and quantities are assumed to have both started in the infinite past.

$$p_{i,t} = p_{i,-\infty} \cdot \ldots \cdot \frac{p_{i,0}}{p_{i,-1}} \cdot \frac{p_{i,1}}{p_{i,0}} \cdot \frac{p_{i,2}}{p_{i,1}} \cdot \ldots \cdot \frac{p_{i,t-1}}{p_{i,t-2}} \cdot \frac{p_{i,t}}{p_{i,t-1}}$$
$$q_{i,t} = q_{i,-\infty} \cdot \ldots \cdot \frac{q_{i,0}}{q_{i,-1}} \cdot \frac{q_{i,1}}{q_{i,0}} \cdot \frac{q_{i,2}}{q_{i,1}} \cdot \ldots \cdot \frac{q_{i,t-1}}{q_{i,t-2}} \cdot \frac{q_{i,t}}{q_{i,t-1}}$$

However, the period-to-period changes are not independently distributed. The sequences are assumed to satisfy a mixing condition, which implies ergodicity; hence, a central limit theorem under weak dependence becomes applicable. Thus, it follows that prices and quantities are marginally log-normally distributed. Having proven marginal log-normal distribution, it follows that they are also jointly log-normally distributed by imposing a functional relationship between prices and quantities and autoregressive relationships within them. $\square$

Proof of Theorem 2.3. The expectation of a log-normally distributed random variable is given by $\exp(\mu + \sigma^2/2)$. After taking natural logarithms it applies that $a \ln X \pm b \ln Y \sim \mathcal{N}(a\mu_X \pm b\mu_Y,\ a^2\sigma_X^2 + b^2\sigma_Y^2 \pm 2ab\sigma_{X,Y})$. Using this and the definitions of the power mean, and the Dutot and unit value indices one finds the following results.

$$P^r = \sqrt[r]{\mathrm{E}\left(\frac{p_{it}^r}{p_{i0}^r}\right)} = \exp\left[\frac{1}{r}\left(r(\mu_{p_t} - \mu_{p_0}) + r^2\frac{\sigma_{p_t}^2 + \sigma_{p_0}^2 - 2\sigma_{p_t,p_0}}{2}\right)\right]$$

$$P^D = \frac{\mathrm{E}(p_{it})}{\mathrm{E}(p_{i0})} = \frac{\exp\left(\mu_{p_t} + \frac{\sigma_{p_t}^2}{2}\right)}{\exp\left(\mu_{p_0} + \frac{\sigma_{p_0}^2}{2}\right)}$$

$$P^{UV} = \frac{\mathrm{E}(p_{it}q_{it})/\mathrm{E}(q_{it})}{\mathrm{E}(p_{i0}q_{i0})/\mathrm{E}(q_{i0})} = \frac{\exp\left(\mu_{p_t} + \mu_{q_t} + \frac{\sigma_{p_t}^2+\sigma_{q_t}^2+2\sigma_{p_t,q_t}}{2}\right) / \exp\left(\mu_{q_t} + \frac{\sigma_{q_t}^2}{2}\right)}{\exp\left(\mu_{p_0} + \mu_{q_0} + \frac{\sigma_{p_0}^2+\sigma_{q_0}^2+2\sigma_{p_0,q_0}}{2}\right) / \exp\left(\mu_{q_0} + \frac{\sigma_{q_0}^2}{2}\right)}$$

By reducing the terms, the proposition follows. □

Proof of Theorem 2.4. Using the definitions of the Laspeyres and Paasche price indices, the expectations are as follows.

$$P^L = \frac{\mathrm{E}(p_{it}q_{i0})}{\mathrm{E}(p_{i0}q_{i0})} = \frac{\exp\left(\mu_{p_t} + \mu_{q_0} + \frac{\sigma_{p_t}^2+\sigma_{q_0}^2+2\sigma_{p_t,q_0}}{2}\right)}{\exp\left(\mu_{p_0} + \mu_{q_0} + \frac{\sigma_{p_0}^2+\sigma_{q_0}^2+2\sigma_{p_0,q_0}}{2}\right)}$$

$$P^P = \frac{\mathrm{E}(p_{it}q_{it})}{\mathrm{E}(p_{i0}q_{it})} = \frac{\exp\left(\mu_{p_t} + \mu_{q_t} + \frac{\sigma_{p_t}^2+\sigma_{q_t}^2+2\sigma_{p_t,q_t}}{2}\right)}{\exp\left(\mu_{p_0} + \mu_{q_t} + \frac{\sigma_{p_0}^2+\sigma_{q_t}^2+2\sigma_{p_0,q_t}}{2}\right)}$$

The proposition follows by reducing the terms. The corresponding power means are found by solving the equations for r. □

Proof of Theorem 2.5. Stationarity in covariance of the processes, i.e. $0 \leq \rho < 1$ and $0 \leq \gamma_1 < 1$, imply that the covariance between any two observations depends only on the lag between them. For the covariance of logarithmic prices, it follows that it is an exponentially decreasing function.

$$\sigma_{p_\kappa,p_\ell} = \gamma_1^{|\kappa-\ell|}\sigma_p^2$$

Using the lag operator and inverting the lag polynom in the function of logarithmic quantities, it can be written as follows.

$$\ln q_{ib} = \alpha + \beta^* \sum_{\tau=0}^{\infty} \rho^\tau \ln p_{ib-\tau} + \left(\sum_{\tau=0}^{\infty} \rho^\tau \mathbf{x_{ib-\tau}} \right) (1-\rho)\boldsymbol{\delta} + \left(\eta_i^q + \sum_{\tau=0}^{\infty} \rho^\tau \varepsilon_{ib-\tau}^q \right)$$

Taking the expectation and subtracting it on both sides yields the following expression.

$$\ln q_{ib} - \mu_q = \beta^* \sum_{\tau=0}^{\infty} \rho^\tau (\ln p_{ib-\tau} - \mu_p) + \sum_{\tau=0}^{\infty} \rho^\tau \varepsilon_{ib-\tau}^q$$

Multiplying this expression with $\ln p_{i\xi} - \mu_p$ and taking the expectation results in the desired covariances.

$$\sigma_{p_\xi, q_b} = \beta^* \sum_{\tau=0}^{\infty} \rho^\tau \sigma_{p_\xi, p_{b-\tau}} = \beta^* \sigma_p^2 \sum_{\tau=0}^{\infty} \rho^\tau \gamma_1^{|\xi-(b-\tau)|}$$

Substituting the appropriate expressions for ξ and b, either 0 or t, the proposition follows by applying the formula for the sum of a geometric series. □

Proof of Theorem 2.6. Substituting the respective expressions into the equations directly yields the stated results. Under the stationarity in covariance assumption, the difference of (co-)variances at different points in time vanishes and approaches zero. For the powers corresponding to the Laspeyres and Paasche price indices, r_L and r_P, respectively, it is assumed that the product of the autoregressive parameters is sufficiently small to be negligible, i.e. the sluggishness of adjustment of quantities or the persistence of the process of prices is low: $\rho\gamma_1 \to 0$. The power corresponding to the Paasche price index is derived under the additional assumptions of sufficiently large t in order for the serial correlation to converge to zero: $\rho^t \to 0$ and $\gamma_1^t \to 0$. □

Proof of Theorem 2.7. The proposition follows directly by reducing the terms. □

Appendix B: CES Price Index

There is a kind of similarity between power means and the CES price index. The latter is derived in Definition 2.4.

Definition 2.4. A constant elasticity of substitution utility function is given by:

$$U(\mathbf{q_b}) = \left(\sum_{i=1}^{n} a_i q_{ib}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \; a_i > 0 \, \forall \, i. \tag{2.32}$$

The CES utility function owes its name to its property of having a constant elasticity of substitution:

$$\frac{\mathrm{d} \ln(q_{jb}/q_{ib})}{\mathrm{d} \ln \left(\frac{\partial U(\mathbf{q_b})}{\partial q_{ib}} / \frac{\partial U(\mathbf{q_b})}{\partial q_{jb}} \right)} = \sigma \, \forall \, i \neq j. \tag{2.33}$$

The associated demand function is found by cost minimisation subject to a constant level of utility:

$$q_{ib}(\mathbf{p_b}, \bar{U}) = \left(\frac{C(\mathbf{p_b}, \bar{U}) \bar{U}^{\frac{1-\sigma}{\sigma}} a_i}{p_{ib}} \right)^{\sigma}. \tag{2.34}$$

This is visualised in Figure 2.7, where the indifference curve (IC) is held fixed while the budget constraint (BC) is shifted parallel until it is tangent to the indifference curve.

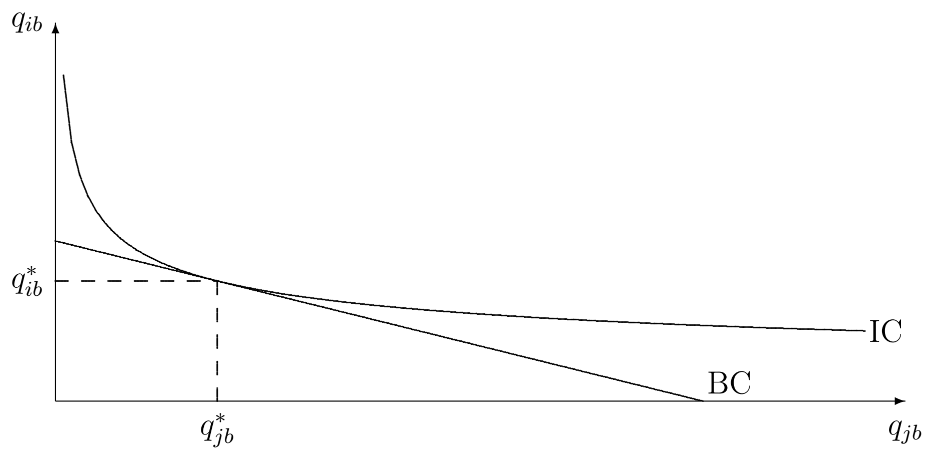


Figure 2.7: Cost Minimisation Subject to a Constant Level of Utility

The demand function follows from the Lagrange multiplier being the marginal cost of utility; moreover, the CES utility function has constant returns to scale, i.e. $U(\phi \mathbf{q_b}) = \phi U(\mathbf{q_b}) \, \forall \, \phi > 0$ and thus, the marginal cost equals the average cost of utility, $\lambda = C(\mathbf{p_b}, \bar{U})/\bar{U}$. The cost function itself is found to be:

$$C(\mathbf{p_b}, \bar{U}) = \sum_{i=1}^{n} p_{ib} q_{ib} = \bar{U} \left(\sum_{i=1}^{n} a_i^{\sigma} p_{ib}^{1-\sigma} \right)^{\frac{1}{1-\sigma}} . \tag{2.35}$$

Lloyd (1975), and Moulton (1996) define the CES price index as a COLI under CES preferences as in Equation (2.32) as follows:

$$P^{LM} = \frac{C(\mathbf{p_t}, \bar{U})}{C(\mathbf{p_0}, \bar{U})} = \left[\sum_{i=1}^{n} \left(\frac{p_{it}}{p_{i0}} \right)^{1-\sigma} \frac{p_{i0} q_{i0}}{\sum_{i=1}^{n} p_{i0} q_{i0}} \right]^{\frac{1}{1-\sigma}} . \tag{2.36}$$

Equation (2.36) is derived by substituting a_i^{σ} from Equation (2.34) in Equation (2.35). Note that for the zero substitutability case, $\sigma = 0$, the Lloyd-Moulton price index becomes the Laspeyres price index, while in the case of $\sigma \to 1$ the geometric (or logarithmic) Laspeyres price index results.

The Lloyd-Moulton price index is an approximation to a true price index which can be calculated with the data at hand for the calculation of the Laspeyres price index plus a minimal extra informational requirement: an estimate of the elasticity of substitution. In this sense, the power mean is the analogue at the elementary level, allowing for the calculation of Laspeyres and Paasche price indices just from information on price relatives plus the price elasticity. However, the notable difference between the Lloyd-Moulton price index and the power mean is that while the former one is a weighted index at the aggregate level, the latter one is an unweighted index at the elementary level. The Lloyd-Moulton price index from Equation (2.36) is yet another aggregate index, notwithstanding its similar construction to the power mean in Equation (2.5). This helps to explain why it is a mean of order $1 - \sigma$ although the price elasticity of demand that is implied by CES preferences is minus the elasticity of substitution, $-\sigma$. Moreover, there is no elasticity of substitution as in Equation (2.33) for the same good at the very elementary level. Eventually, the Lloyd-Moulton price index faces the same problem as any other aggregate index: "which unweighted index formula should be used at the elementary level, where no expenditure share weights are available?"

Appendix C: Numerical Example

Consider the following example of $K = 2$ groups of goods with $n_k = 2$ goods each and their respective prices and quantities (in grey).

k	j_k	p_0	q_0	p_t	q_t	p_t/p_0	$p_0 \cdot q_0$	$p_t \cdot q_t$
1	1	7	3	14	1	2.00	21	14
	2	20	10	22	7	1.10	200	154
	Σ		13		8		221	168
2	1	15	25	60	13	4.00	375	780
	2	2	63	17	36	8.50	126	612
	Σ		88		49		501	1,392

This represents the outcome of the true but unknown data generating process. In price statistics one can only observe price relatives for each good in a group of goods and aggregate base period expenditures (more precisely, the share weights) at the group of goods level (in black). In this case the Laspeyres price indices take on the values $P_1^L = 1.19$ and $P_2^L = 5.13$ which correspond to power means with $r_1 = -9.20$ and $r_2 = -1.95$, respectively. The Carli indices ($r = 1$) are $P_1^C = 1.55$ and $P_2^C = 6.25$. Thus, the two-staged Laspeyres price index with Carli indices at the elementary level becomes $P^{L(C)} = 4.81$, well away from its true value of $P^L = 3.92$.

Chapter 3

Sources of Revisions of Seasonally Adjusted Real Time Data

3.1 Introduction

The importance of real time data becomes obvious when one tries to understand economic policy decisions that were made based on historical data and reconsiders these past situations in the light of more recent data (Orphanides, 2001). Statistical agencies and users of seasonally adjusted real time data alike are interested in it, inter alia in terms of the quality and interpretation of statistics. Statistical agencies are concerned about the quality and usefulness of the statistics they publish and how to increase these. Users of seasonally adjusted real time data need to know the extent of revisions for economic analysis and forecasts and how to take them into account. Thus, revisions of seasonally adjusted real time data are a frequently discussed topic, while their sources are often disregarded (Damia and Picón Aguilar, 2006).

This chapter deals with the empirical quantification of these sources for selected German time series, using the real time database of the Deutsche Bundesbank. The investigated time series are important business cycle indicators: (1) real gross domestic product, (2) employment, (3) output in and (4) orders received by the manufacturing sector as well as (5) retail trade turnover. Generally, revisions to seasonally adjusted real time data have two separate but inter-related sources. One of these sources is the technical procedure of the method used for seasonal adjustment, in Germany Census X-12-ARIMA: the release of new unadjusted data, old unadjusted data remaining unchanged, leads to a shift of the base period and a change in the weights of smoothing filters. The other is the revision process of unadjusted data in real time: on its first date of release the data contains estimates for missing values which will be updated by and by with actual figures.

The contribution to the literature is to empirically quantify the uncertainty of seasonally adjusted real time data in terms of revisions and decompose them into these two sources. The revision process depends heavily on the properties of the time series.

The outline is as follows. The next section gives definitions, Section 3.3 a description of the methodology. A model for the decomposition of revisions is developed in Section 3.4. In Section 3.5 the results, i.e. actual revisions and their decomposition, are presented. The final section concludes.

3.2 Definitions

The development of seasonal adjustment traces back to the knowledge of astronomy and meteorology in the early 19^{th} century. The decomposition of a time series into unobservable components was first applied in these sciences. This still remains the core of modern seasonal adjustment. It is assumed that the unadjusted time series u_t is decomposed into a trend-cycle component c_t, a seasonal component s_t (assume, for the sake of simplicity, that calendar effects are included in the seasonal component) and an irregular component i_t. Formally, the decomposition of the multiplicative model, which is the most important in practice, can be written as

$$u_t = c_t \cdot s_t \cdot i_t. \tag{3.1}$$

The aim of seasonal adjustment is to calculate the seasonally adjusted time series a_t.

$$a_t := \frac{u_t}{s_t} = c_t \cdot i_t \tag{3.2}$$

Its relative period-to-period changes in per cent are denoted Δ_t.

$$\Delta_t := \frac{a_t}{a_{t-1}} - 1 \tag{3.3}$$

Let $a_{t|t}$ denote the first estimate of the seasonally adjusted time series as given in Equation (3.2) at time t based on unadjusted data up to time t and let $a_{t|T}$ be its most recent estimate at time t based on unadjusted data up to time T, $t < T$. Then the per cent revision of the seasonally adjusted time series r_t^a is defined as the relative deviation of the most recent estimate from the first one. This can be interpreted as the answer to the question of how much a given adjustment is affected by appending new and updating old unadjusted data.

$$r_t^a := \frac{a_{t|T}}{a_{t|t}} - 1 \approx \ln \frac{a_{t|T}}{a_{t|t}}. \tag{3.4}$$

This could approximately be reformulated as

$$r_t^a \approx \ln \frac{u_{t|T}}{u_{t|t}} - \ln \frac{s_{t|T}}{s_{t|t}} \approx r_t^u - r_t^s \tag{3.5}$$

which follows directly from Equations (3.2) and (3.4) and is the difference between revisions of the unadjusted time series and those of the seasonal component.

Revisions of per cent period-to-period changes r_t^Δ, measured in percentage points, are defined as

$$r_t^\Delta = \Delta_{t|T} - \Delta_{t|t}, \tag{3.6}$$

where $\Delta_{t|T}$ and $\Delta_{t|t}$ are defined analogously to Δ_t in Equation (3.3).

3.3 Methodology

The seasonal component accounts for regular patterns arising from annually recurring circumstances, like seasons and length of months. These regular patterns contain no news for analysing the current economic development. Thus, the intention of seasonal adjustment is to eliminate these patterns from the unadjusted time series, whilst maintaining news in the seasonally adjusted time series (cf. Eurostat, 2009). Eventually, period-to-period changes of the seasonally adjusted time series are to be interpreted economically. For a description of seasonal adjustment with the X-12-ARIMA method see Appendix A.

As the seasonal adjustment method X-12-ARIMA is not model based, it is not possible to calculate standard errors for the seasonally adjusted time series straightforwardly. The scientific literature on this topic is limited (for references see Eurostat, 2006). The existing approaches often share the assumption of pure linear smoothing. However, logarithmic transformation, RegARIMA modelling and outlier detection, all three commonly used in practice, are neither linear nor can they be approximated. Due to the fact that X-12-ARIMA lacks statistical properties, an ad hoc analysis will be presented here.

Table 3.1 shows the most important basic data for the seasonal adjustment of the analysed time series (for a detailed description of real time data properties see Appendix B). Common to all time series within this chapter is the period covered from the beginning of 1991 to the end of 2006. The forecast horizon of RegARIMA modelling is twelve periods in each case, all time series are seasonally adjusted assuming the multiplicative model.

Table 3.1: Basic Data of Seasonal Adjustment of Time Series

Time Series	Measuring Unit (Base Year)	Calculation Method	Data Frequency
Gross Domestic Product	Index (2000 = 100)	Flow	Quarterly
Employment	1,000 Persons	Stock	Monthly
Output	Index (2000 = 100)	Flow	Monthly
Orders Received	Index (2000 = 100)	Flow	Monthly
Retail Trade Turnover	Index (2003 = 100)	Flow	Monthly

Analysis of revisions is based on the six-year period from 1996 to 2001. The five years at the beginning of the data set are not used in revision analysis to estimate primary seasonally adjusted figures of acceptable quality, and a five-year period at the end is left in order to compare the results with almost final seasonally adjusted figures. The real time database of the Deutsche Bundesbank allows the impact of revisions of real time data to be quantified. However, as the real time database is currently under construction, seasonally adjusted real time data are not yet available over a long data span. Therefore, the analysis is performed on the basis of selected, manually extended time series starting in 1996 (Gerberding et al., 2005).

3.4 Model

3.4.1 Seasonal Adjustment Basics

Equation (3.2) for the seasonally adjusted time series illustrates that generally revisions to seasonally adjusted real time data a_t have two separate but inter-related sources. One source is the technical procedure of the method used for seasonal adjustment, responsible for s_t, the other is the revision process of unadjusted data in real time, u_t. These sources are broken down in Figure 3.1. The estimation of the seasonally adjusted time series a_t depends on the estimate of the seasonal component s_t and the quality of unadjusted data u_t. In turn, the estimate of the seasonal component is influenced by the quality of unadjusted data.

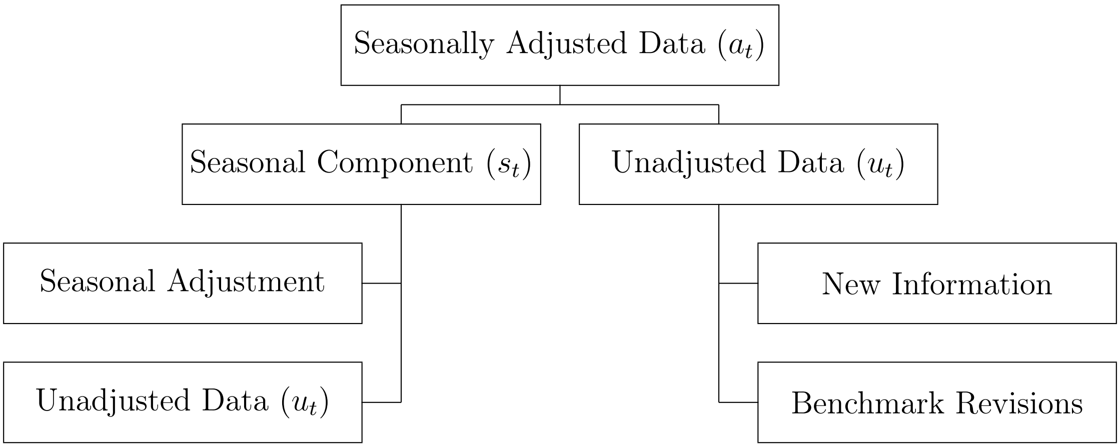


Figure 3.1: Sources of Revisions

In other words, revisions of unadjusted data influence the estimation of the seasonal component. This can also be seen in Figure 3.2, which shows the structure of the X-12-ARIMA method (Deutsche Bundesbank, 1999). RegARIMA modelling constitutes a major improvement in X-12-ARIMA compared to its predecessor version X-11. The capabilities of RegARIMA modelling are calendar adjustment and the extension of a time series beyond its end. Forecasted time series values are used for applying symmetric smoothing filters in the seasonal adjustment core and forecasting seasonal components. The seasonal adjustment core represents the established and enhanced X-11 method. This method is an iterative, mathematical procedure for seasonally adjusting time series.

Table 3.2: Data Structure for Revision Analysis†

	Seasonal Adjustment				Real Time Data			
t	1	2	...	T	1	2	...	T
1	$a^1_{1\|T}$				$a_{1\|1}$			
2	$a^2_{1\|T}$	$a^2_{2\|T}$			$a_{1\|2}$	$a_{2\|2}$		
$\vdots$	$\vdots$	$\vdots$	$\ddots$		$\vdots$	$\vdots$	$\ddots$	
T	$a_{1\|T}$	$a_{2\|T}$	...	$a_{T\|T}$	$a_{1\|T}$	$a_{2\|T}$	...	$a_{T\|T}$

† Reporting dates in columns and dates of releases in rows.

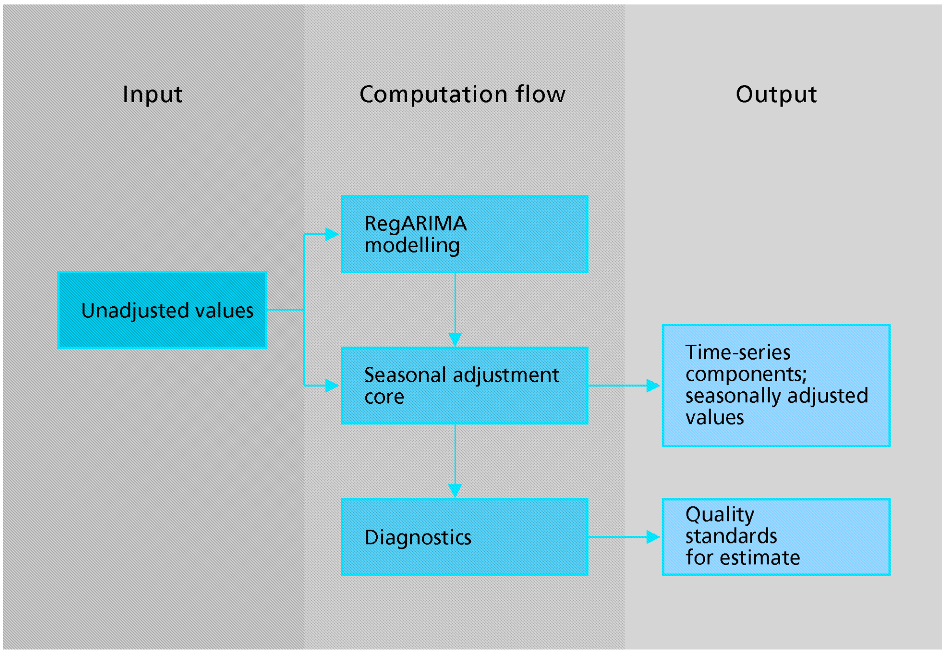


Figure 3.2: Structure of X-12-ARIMA

The data structure for revision analysis is illustrated in Table 3.2. For the analysis of the non-real time component, that is the seasonal adjustment method, the most recent time series, truncated at the respective time t which is indicated by superscript t, is used, while for the analysis of real time data the data actually available at that specific point in time are taken.

In order to separate both effects a pure seasonal adjustment revision $\tilde{r}_t^s$ is calculated which is not related to the revision of unadjusted data r_t^u. The latter is defined analogously to r_t^a in Equation (3.4): $r_t^u := u_{t|T}/u_{t|t} - 1$. The former has to take into consideration the non-real time data component. Thus, $\tilde{r}_t^s$ does not measure revisions of the seasonal component r_t^s in real time but those of the seasonally adjusted time series sequentially, much like the automatic *History* procedure implemented in X-12-ARIMA (U. S. Census Bureau, 2001). This is similar to r_t^a with the difference that unadjusted data are not updated but the most recent time series is truncated at time t and hence revisions of unadjusted real time data do not influence seasonal adjustment revisions, which are per definitionem non-real time. The definition of $\tilde{r}_t^s$ is equivalent to r_t^a, where the denominator $a_{t|T}^t$ is now defined as the corresponding element of the truncated time series. Analogous definitions apply to period-to-period changes.

Hence, revisions of seasonally adjusted real time data, r_t^a, comprise two effects that show up in revisions resulting from seasonal adjustment, $\tilde{r}_t^s$, and those from unadjusted real time data, r_t^u. A simple approach to the decomposition of revisions is to understand r_t^s but not $\tilde{r}_t^s$ as the residual of the approximate equality in Equation (3.5) (for references see OECD, 2007).

$$r_t^s \approx r_t^u - r_t^a \neq \tilde{r}_t^s \tag{3.7}$$

It is important to note that r_t^s and $\tilde{r}_t^s$ are two fundamentally different measures. r_t^s is defined to be the residual of a one-to-one relationship and is influenced by another source of revisions. On the contrary, $\tilde{r}_t^s$ measures a pure seasonal adjustment effect and allows for an original decomposition of total revisions into their sources. Thus, in the remainder, the focus is exclusively on the pure seasonal adjustment revision $\tilde{r}_t^s$ and not on revisions of the seasonal component r_t^s.

3.4.2 Data and Theoretical Background

Data used in this study are firstly unadjusted real time data rebased to the current base year but without metadata for filtering benchmark revisions, and secondly the present user setting of seasonal adjustment which is held constant throughout all vintages. For that, a new procedure is developed and implemented within the framework of the seasonal adjustment method X-12-ARIMA in this chapter. It measures revisions of real time data, each time rerunning seasonal adjustment with the latest information available. RegARIMA model parameters, namely calender regressors, outliers and ARMA parameters, are estimated with the full data span of the latest time series available. In doing so a third source of revisions is suppressed that emerges if user settings of seasonal adjustment are changed. Therefore, the revisions derived from seasonal adjustment mainly describe the properties of seasonal filters and extreme value detection. For seasonal adjustment purposes, data prior to 1996 are added to each vintage using the latest time series available. Using the two aforementioned data sources, the real time data is seasonally readjusted, i.e. historical published figures are not used.

The aim is to decompose total revisions, those of seasonally adjusted real time data, into a seasonal adjustment and an unadjusted real time data part. The hypothesis is that the variance of revisions of seasonally adjusted real time data is the result of revisions from these two sources. To reiterate, the former source results from shifting smoothing filters weights, the latter from newly available information. Basically, a causal relationship of the form $r_t^a = f(\tilde{r}_t^s, r_t^u)$ is presumed. Revisions can be considered in a panel regression model, taking into account time series and reporting date specific effects. The model chosen is a two-way fixed effects heterogeneous panel regression model. Random effects are infeasible in an unbalanced sample and the within-estimator is asymptotically unbiased. The formal model is stated in Equation (3.8).

$$r_{i,t}^a = \alpha_{i,t} + \beta_i^s \cdot \tilde{r}_{i,t}^s + \beta_i^u \cdot r_{i,t}^u + \nu_{i,t},\ \alpha_{i,t} = \alpha + \gamma_i + \delta_t \tag{3.8}$$

Different time series are indicated by subscript i, γ_i and δ_t are dummy variables accounting for the heterogeneity across time series and reporting dates, $\nu_{i,t}$ is assumed to be white noise. Slope coefficients β_i^s and β_i^u are allowed to vary across time series to capture their unique properties. In the case of real gross domestic product, figures are assigned to the month in which the quarter ends.

Estimated slope coefficients represent marginal effects and hence depend on the order of magnitude of the underlying revisions. They could be used to calculate curve elasticities, ε_i^s and ε_i^u, employing average absolute revisions, $\bar{R}_i^a$, $\bar{\tilde{R}}_i^s$ and $\bar{R}_i^u$. For this, absolute values are used, since otherwise revisions would cancel out to a large extent. Elasticities have the advantage that their value possesses a direct interpretation: values less than one indicate inelasticity, values greater than one elasticity. The formulae for the calculation of curve elasticities are specified in Equation (3.9). These formulae are easily derived from the definition of elasticity as the relative change of the endogenous variable y to a 1% change of the exogenous variable x: $\varepsilon := \mathrm{d}\ln y / \mathrm{d}\ln x = (\mathrm{d}y/\mathrm{d}x) \cdot (x/y)$.

$$\varepsilon_i^s := \beta_i^s \cdot \frac{\bar{\tilde{R}}_i^s}{\bar{R}_i^a} \qquad \varepsilon_i^u := \beta_i^u \cdot \frac{\bar{R}_i^u}{\bar{R}_i^a} \tag{3.9}$$

The following statistical tests have been carried out on the estimated model. Coefficients are tested with a one-sided t-test for significance of the null hypothesis $\beta_i = 0$ against the alternative $\beta_i > 0$. The same test strategy is applied to elasticities with the alternative hypothesis of inelastic revisions, i.e. $\varepsilon_i = 1$ against $\varepsilon_i < 1$. For the ratio of elasticities, the null hypothesis $\varepsilon_i^u/\varepsilon_i^s = 1$ is used against the alternative $\varepsilon_i^u/\varepsilon_i^s > 1$, which means that r_t^u has a larger effect on r_t^a than $\tilde{r}_t^s$. Standard errors of elasticities and their ratio are calculated using the Delta Method. Significance levels are derived assuming asymptotic normality. χ^2-Wald tests are employed to test for parameter homogeneity, i.e. testing $\beta_i = \beta_j \,\forall\, i \neq j$ under the null hypothesis. An F-test is suitable for checking model adequacy, the null hypothesis is redundant fixed effects.

A pseudocode for the calculation of seasonally adjusted figures and elasticities as well as the description of the Delta Method for the estimation of standard errors of the latter are to be found in Appendix C.

3.5 Results

3.5.1 Predictability and Revisions

Certainty of seasonal adjustment at the end of the time series depends on the forecast quality of RegARIMA modelling. With the aid of RegARIMA modelling a time series is extended beyond its end and symmetric or less asymmetric smoothing filters become applicable. Generally smaller revisions result from extending the time series rather than applying asymmetric smoothing filters (Kirchner, 1999). The closer the forecast comes to the realised value, the smaller are the revisions. On the other hand, extending a time series which is hard to forecast will worsen the robustness of the results. The average absolute percent ex ante-forecast errors over the last three years are visualised in Figure 3.3. While time series of real gross domestic product and retail trade turnover can be relatively well predicted, predictability of time series of output in and orders received by the manufacturing sector is significantly worse. The time series of employment performs best in these terms.

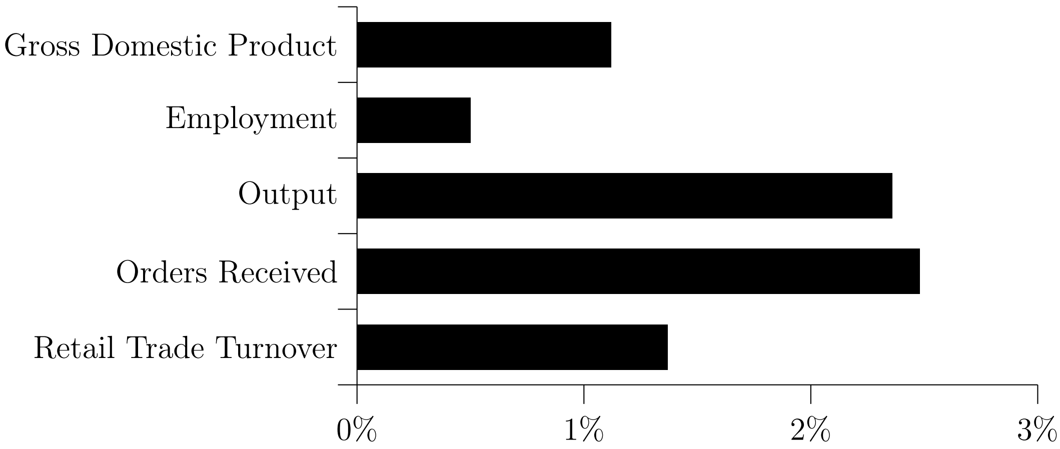


Figure 3.3: Average Absolute Percent Ex Ante-Forecast Errors

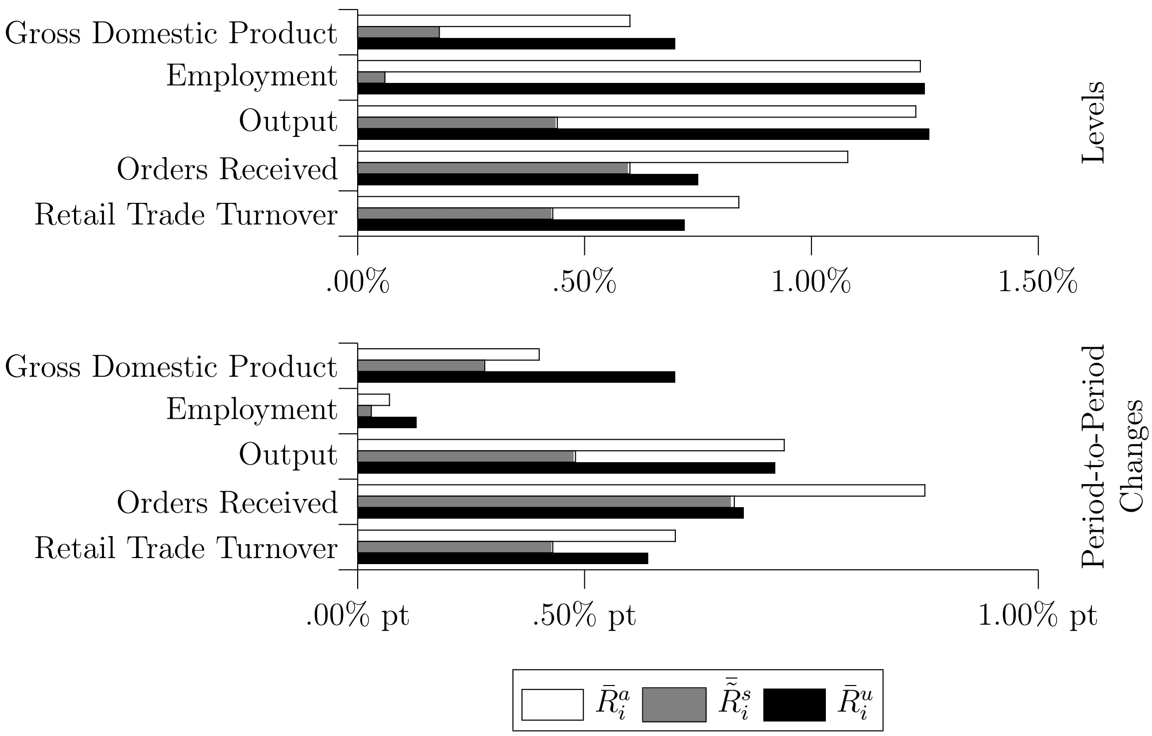


Figure 3.4: Average Absolute Revisions

Time series specific average absolute revisions of seasonally adjusted real time data, seasonal adjustment and unadjusted real time data and their corresponding period-to-period changes are drawn in Figure 3.4. It is conspicuous that revisions of seasonally adjusted real time data are much higher than those arising from seasonal adjustment both for levels and period-to-period changes. Unadjusted real time data revisions are by and large approximately of the same magnitude as those of seasonally adjusted real time data. The problem with revisions of real time data levels of employment is that the vintages exhibit a major structural change due to the labour market reform and the inclusion of the new concept of persons in marginal employment: revisions for the period before 1999 are a small fraction of those for the period after 1999. Revisions of period-to-period changes of orders received by the manufacturing sector are very high due to the occurrence of large orders. Extreme value replacement for these outliers in the seasonal adjustment core depends on the vintages.

3.5.2 Time Series Properties

A large influence on the magnitude of revisions and for this reason on the uncertainty of seasonal adjustment could be ascribed to two factors. One of them is the standard deviation of the irregular component which describes the time series' fluctuations and the problems of forecasting and seasonal adjustment, see also Figure 3.3. The other one is the standard deviation of the seasonal component which represents the dependence of the time series on seasonal impacts and the degree of necessity of seasonal adjustment. In the extreme case of no seasonal fluctuations the unadjusted time series is equal to the seasonally adjusted time series, $u_t = c_t \cdot i_t = a_t$.

For comparison these influences are illustrated in Figure 3.5 as percentages. Dividing the standard deviation by the expectation yields the dimensionless coefficient of variation. Here, the expectation is equal to one in both cases.

The two components interact during seasonal adjustment (cf. Appendix A for the significance of the seasonal-irregular component). Therefore, technically the difficulties remain rather small if the standard deviation of the seasonal component is relatively low, even if the one of the irregular component is relatively high. The

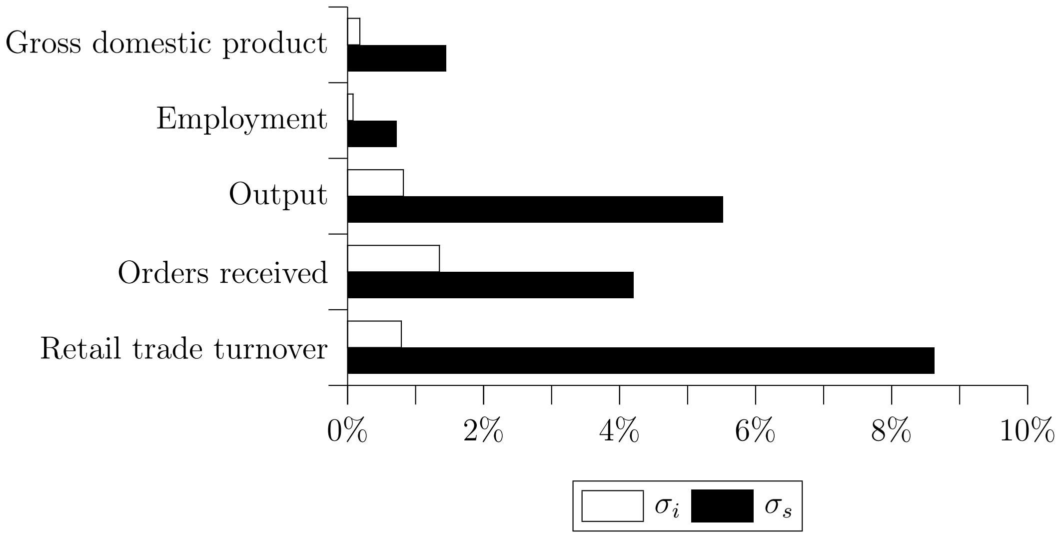


Figure 3.5: Standard Deviations of the Irregular and the Seasonal Component

time series of real gross domestic product and employment have small standard deviations of both the irregular and the seasonal component, whereby seasonal adjustment is unproblematic. The biggest difficulties arise with the time series of output in and orders received by the manufacturing sector and retail trade turnover. These time series have high standard deviations of the irregular as well as the seasonal component which leads to larger revisions.

With these five observations a cross-section regression is carried out with the average absolute revisions of seasonal adjustment (cf. Figure 3.4) on the standard deviations of the irregular and the seasonal component. The result with a coefficient of determination of $R^2 = 0.95$ is stated in Equation (3.10) with Newey-West standard errors in brackets beneath the parameters.

$$\hat{\hat{\tilde{R}}}_i^s = \underset{(0.0130)}{0.4044} \cdot \sigma_i + \underset{(0.0042)}{0.0156} \cdot \sigma_s \tag{3.10}$$

It follows that the standard deviations of the irregular and the seasonal component both influence the revisions positively. The higher the irregularity and the seasonality of a time series are, the higher are the revisions. The standard deviation of the irregular component has a higher impact on the revisions than the one of the seasonal component.

3.5.3 X-11 Statistics

Whether or not a change of the seasonally adjusted time series is random is a vital question. The Deutsche Bundesbank solves this problem in short-term business cycle analysis, e.g. in the Monthly Reports, by putting two or three periods at the end of a time series together and comparing their change to an equally long period before. Here, two diagnostics are presented which will give an indication and support for answering the questions whether a change of the seasonally adjusted time series is significant and how many periods should be put together in short-term business cycle analysis. In any case, they will not provide an exact solution to the problem. The diagnostics are the average duration of growth and decline phases of the seasonally adjusted time series and the number of periods until the ratio of the irregular and the trend-cycle component becomes less than one.

The average duration of growth and decline phases of the seasonally adjusted time series (average duration of run, ADR) is the average number of successive positive and negative terms, respectively, of the period-to-period changes. If a term is zero, it is counted in the current phase. This is a suitable historical indicator. The number of periods until the ratio of the irregular and the trend-cycle component becomes less than one (periods for cyclical dominance, PCD) shows when the trend-cycle component dominates the development of the seasonally adjusted time series and not the irregular component. This relationship comes from the definition of the seasonally adjusted time series as $c_t \cdot i_t$ and the definition of the irregular/trend-cycle component ratio as $\bar{i}/\bar{c}$.

Figure 3.6 visualises the average duration of run and the periods for cyclical dominance. While PCD can only take integer values, ADR as an average is not restricted to be an integer period.

Large values for average duration of run coincide with small values for periods for cyclical dominance. The coefficient of correlation between these two indicators is -0.65.

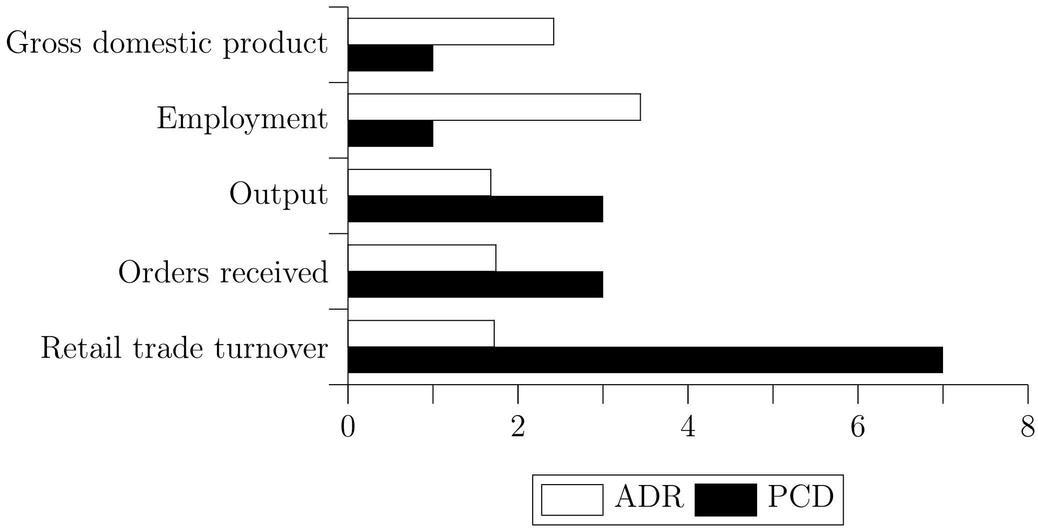


Figure 3.6: X-11 Diagnostics

3.5.4 Decomposition of Revisions

The model in Equation (3.8) was estimated and the results are presented here. Besides estimated slope coefficients, elasticities as in Equation (3.9) and their ratio are calculated and reported below. All estimated slope coefficients are greater than zero on the 1% significance level with the notable exception of period-to-period changes of employment.

Results for levels in Table 3.3 reveal that elasticities of seasonal adjustment revisions are significantly less than one, i.e. revisions of seasonally adjusted real time data react on average relatively inelastically to revisions of seasonal adjustment. Revisions of unadjusted real time data are somewhat mixed, the null hypothesis cannot be rejected for real gross domestic product and is only weakly rejected for employment and output in the manufacturing sector. Ratios of these elasticities clearly show that revisions of unadjusted real time data and not those of seasonal adjustment are the major source of seasonally adjusted real time data revisions. They are 1.25 to 2.81 times larger than the latter; the value for employment should not be overrated as its standard error is extremely large.

Table 3.3: Regression Results for Levels†

Time Series	β_i^s	β_i^u	ε_i^s	ε_i^u	$\varepsilon_i^u/\varepsilon_i^s$
Gross Domestic Product	1.50***	.72***	.46***	.85	1.86**
Employment	1.04***	.91***	.05***	.91*	17.46***
Output	.93***	.90***	.33***	.92**	2.81***
Orders Received	.98***	.98***	.55***	.69***	1.25*
Retail Trade Turnover	.96***	.80***	.49***	.69***	1.40**

† Observations = 260, R^2 = .99, $\chi_s^2(4) = 12.61^{**}$, $\chi_u^2(4) = 35.63^{***}$, $\alpha = .00$, $F(73, 176) = 1.48^{**}$. p-values: *** = 1%, ** = 5%, * = 10%.

For period-to-period changes the results in Table 3.4 are not that clear cut. Results for employment are insignificant due to the aforementioned problem. Elasticities of revisions of real gross domestic product are not significantly less than one, whereas those of the other time series are. Except for the ratio of elasticities of output in the manufacturing sector, which is found to be significantly greater than one, ratios are not statistically different from one as the inversion of the null hypothesis does not hold either.

Table 3.4: Regression Results for Period-to-Period Changes†

Time Series	β_i^s	β_i^u	ε_i^s	ε_i^u	$\varepsilon_i^u/\varepsilon_i^s$
Gross Domestic Product	1.33***	.45***	.95	.78	.83
Employment	.78	.12	.34**	.23***	.65
Output	.93***	.72***	.48***	.71***	1.49***
Orders Received	.98***	.87***	.65***	.59***	.91
Retail Trade Turnover	.91***	.68***	.56***	.62***	1.11

† Observations = 260, R^2 = .99, $\chi_s^2(4) = 14.44^{***}$, $\chi_u^2(4) = 96.33^{***}$, $\alpha = -.03^{***}$, $F(73, 176) = 1.98^{***}$. p-values: *** = 1%, ** = 5%, * = 10%.

Coefficients of determination are high for both models at $R^2 = 0.99$. χ^2-Wald tests indicate that the heterogeneous model is appropriate as the parameter homogeneity hypothesis is rejected. In line with this are F-tests which support the adequacy of the fixed effects model.

Although results for levels clearly indicate the importance of unadjusted real time data revisions and those for period-to-period changes do not contradict them, it is worth taking a closer look at the latter because these are important for short-term business cycle analysis. Due to the uncertainty of seasonally adjusted real time data, at the end of the time series a two or three-period moving average is often used in practice and its change is compared to its previous period counterpart. This strategy is used here. Some time series, such as employment, have a small standard error, some a much higher one, such as output in and orders received by the manufacturing sector. This is the only time series where the standard deviation of revisions of seasonal adjustment is greater than that of unadjusted real time data. By calculating moving averages, the standard errors are lowered as noise is partially smoothed out.

The results in Table 3.5 show that in general ratios of elasticities become greater than one when using moving averages (MA). However, results mainly remain insignificant. This is not true for orders received by the manufacturing sector whose ratio becomes significantly less than one. A possible explanation is the placement of large orders which is done more sporadically than periodically. This problem needs further investigation before robust estimates can be calculated, as even in this case revisions of period-to-period changes of unadjusted real time data are greater than those of seasonal adjustment. Additionally, the likelihood of estimating the wrong sign of period-to-period changes decreases. Thus, revisions of unadjusted real time data become more important since their elasticity increases absolutely and relatively, and revisions of period-to-period changes themselves do not have such a large influence on short-term business cycle analysis as the sign does not change extraordinarily often.

Table 3.5: Ratios of Elasticities of Period-to-Period Changes of Moving Averages

Time Series	Original	MA(2)	MA(3)
Gross Domestic Product	.83	1.95**	1.17
Employment	.65	.97	1.05
Output	1.49***	1.17	1.21
Orders Received	.91	.72	.71
Retail Trade Turnover	1.11	1.18	1.09

3.6 Summary

This chapter deals with the decomposition of sources of revisions of seasonally adjusted real time data. On the one hand, non-real time data revisions of the seasonal adjustment method as a result of appending new unadjusted data are discussed. On the other hand, the focus is on revisions of unadjusted real time data, resulting from updating old unadjusted data. The framework for decomposition of total revisions into these two sources is a heterogeneous panel regression model. As estimated slope coefficients indicate marginal effects, elasticities are calculated and set in relation to each other. It can be concluded that revisions of unadjusted real time data play an important role when trying to understand the revision process of seasonally adjusted real time data. For levels, elasticities of seasonally adjusted real time data revisions with respect to unadjusted real time data revisions are greater than those with respect to revisions stemming from the seasonal adjustment method. They are 1.25 to 2.81 times larger than the latter. For period-to-period changes results are somewhat mixed. However, the calculation of period-to-period changes based on moving averages underlines the importance of unadjusted real time data revisions for explaining those of seasonally adjusted real time data. Ratios of elasticities become greater than one. Furthermore, this chapter confirms a well-known result for the recent past. The current domain of uncertainty of seasonal adjustment depends heavily on the time series analysed and their properties. Nearly throughout this chapter the same time series perform equally well or worse.

Refinements of decompositions applied here might be to augment the model to account for benchmark revisions and to incorporate metadata, both not the focus of this chapter (Knetsch and Reimers, 2009). The importance of these factors in practice makes them two further topics for fundamental research. However, rebasing indices is to be preferred over recalculating them, as the new base year does not apply to old data. Another fundamental question relates to revisions due to changes of the user settings of seasonal adjustment, like RegARIMA modelling and length of seasonal filters.

Appendix A: X-12-ARIMA

In what follows RegARIMA modelling in the X-12-ARIMA method and its seasonal adjustment core are explained in detail (cf. Figure 3.2). The illustration here follows the expositions of Findley et al. (1998), and Ladiray and Quenneville (2001), respectively. A new development as regards calendar adjustment in X-12-ARIMA is RegARIMA modelling. This is the first part of the seasonal adjustment procedure. The seasonal adjustment core itself is made up of three parts. The iterative smoothing procedure, contained in sections B, C and D, is displayed in Figure 3.7 (cf. Kirchner, 1999).

RegARIMA Modelling

RegARIMA modelling contains a regression component of natural logarithms of the unadjusted time series $u_t = c_t \cdot s_t \cdot i_t$ (table A 1) according to Equation (3.1):

$$\ln u_t = \boldsymbol{x_t}'\boldsymbol{\beta} + v_t \tag{3.11}$$

and an ARIMA component for the error term $v_t = \ln u_t - \boldsymbol{x_t}'\boldsymbol{\beta}$:

$$\underbrace{\phi(\mathrm{L})}_{\mathrm{AR}(p)}\,\underbrace{\Phi(\mathrm{L}^s)}_{\mathrm{SAR}(P)}\,\underbrace{(1-\mathrm{L})^d}_{\mathrm{I}(d)}\,\underbrace{(1-\mathrm{L}^s)^D}_{\mathrm{SI}(D)}\,v_t = \underbrace{\theta(\mathrm{L})}_{\mathrm{MA}(q)}\,\underbrace{\Theta(\mathrm{L}^s)}_{\mathrm{SMA}(Q)}\,e_t. \tag{3.12}$$

The main regressors in Equation (3.11) are calendar regressors which normally represent the variation in the number of working days and holidays. Furthermore, exogenous regressors are used to model outliers which would otherwise distort the estimation of the model parameters. The ARIMA model in Equation (3.12) can be written as $\ln(p, d, q)(P, D, Q)_s$. L denotes the lag (or backshift) operator ($\mathrm{L}^\tau z_t = z_{t-\tau}$), s the seasonal frequency ($s = 12$ for monthly data, $s = 4$ for quarterly data), $\phi(\mathrm{L}) = (1 - \phi_1\mathrm{L} - \cdots - \phi_p\mathrm{L}^p)$ the non-seasonal AR operator AR(p), $\Phi(\mathrm{L}^s) = (1 - \Phi_1\mathrm{L}^s - \cdots - \Phi_P\mathrm{L}^{Ps})$ the seasonal AR operator SAR(P), $(1-\mathrm{L})^d$ the non-seasonal difference operator I(d), $(1-\mathrm{L}^s)^D$ the seasonal difference operator SI(D), $\theta(\mathrm{L}) = (1-\theta_1\mathrm{L}-\cdots-\theta_q\mathrm{L}^q)$ the non-seasonal MA operator MA(q) and $\Theta(\mathrm{L}^s) = (1 - \Theta_1\mathrm{L}^s - \cdots - \Theta_Q\mathrm{L}^{Qs})$ the seasonal MA operator SMA(Q). For non-seasonal and seasonal AR and MA operators, it is assumed that all roots of the inverse lag polynomial lie inside the unit circle. The error term e_t is assumed to be white noise.

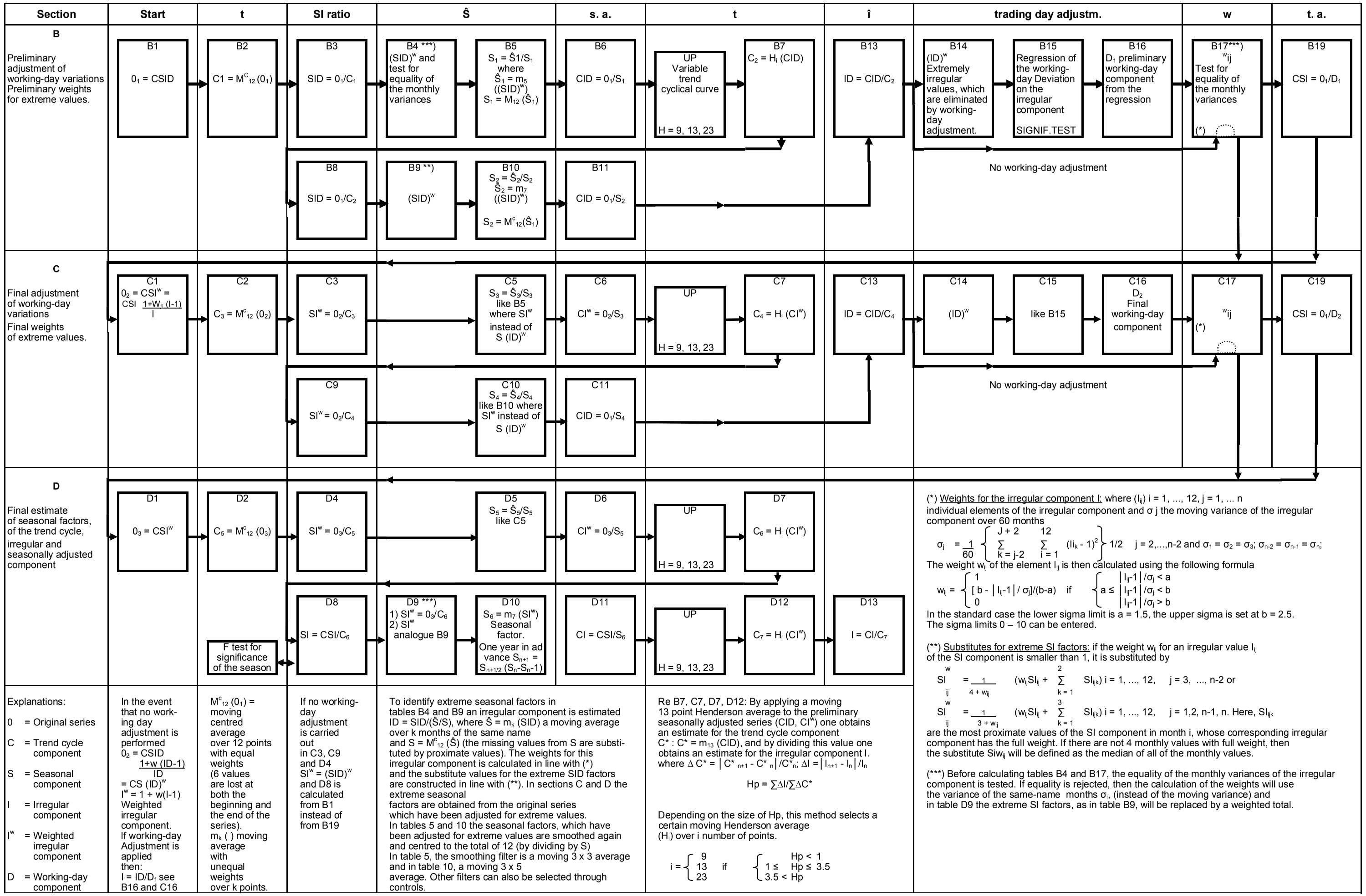


Figure 3.7: Workflow Diagram of the X-11 Seasonal Adjustment Core

The correlogram in Figure 3.8 shows the theoretical autocorrelation structure of a $\ln(0,1,1)(0,1,1)_{12}$ model – here in non-seasonal and seasonal differences – with the coefficients $\theta = .3$ and $\Theta = .6$. This so-called airline model is most frequently chosen, see also Table 3.6. The autocorrelation is evident around the seasonal frequency. As well as the negative autocorrelation at lag $\tau = 1$ and $\tau = 12$, which results directly from the MA coefficients, this model has a positive autocorrelation at lags $\tau = 11$ and $\tau = 13$, which is due to the interaction of the two coefficients. This illustration is the basis for identifying a suitable ARIMA model for practical seasonal adjustment using X-12-ARIMA.

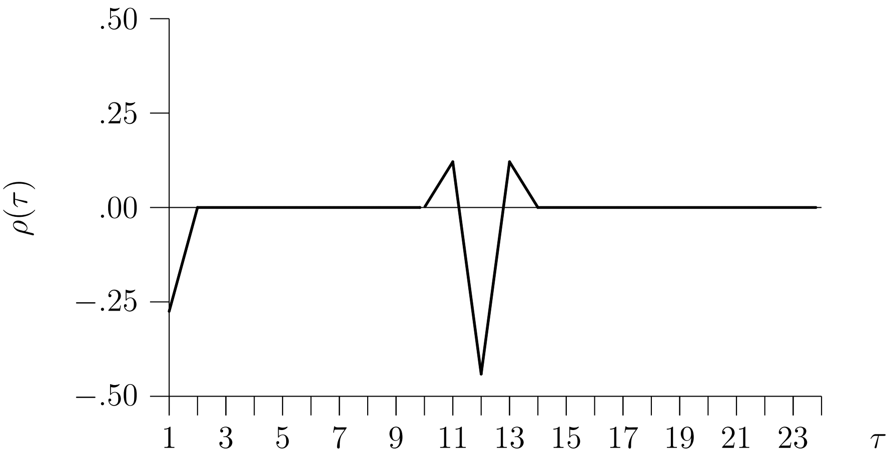


Figure 3.8: Autocorrelation Function (ACF) of the Differenced Airline Model

X-12-ARIMA allows for the adjustment of working-day fluctuations using RegARIMA modelling. To this end, Equations (3.11) and (3.12) are estimated with suitable regressors and the estimation of the calendar component is determined based on the estimated coefficients and the regressor values (table A 9). The calendar component does the same as the seasonal component (cf. Section 3.3 for an explanation of the seasonal component) for recurring calendar events which shift between months, such as the number of working days within the same month in different years and timing of holidays. As the calendar component contains no news on the underlying economic development either, it can be interpreted as seasonality in the broader sense.

Seasonal Adjustment Core

The procedure in the seasonal adjustment core can be simplified into four steps. In the first step, the calendar-adjusted time series (table B 1) is smoothed using a moving average in order to estimate the provisional trend-cycle component. The trend-cycle component c_t mirrors the long-run economic development and business cycle fluctuations, and so is a smooth, generally non-stationary time series. From a theoretical standpoint this component would be the most desirable one for business cycle analysis as it shows cyclical turning points most clearly. However, at the end of the time series, the estimation of this component is not possible with high accuracy. There is a persistent risk of spurious extrapolation of a past trend-cycle into the future. Nonetheless, an estimate is needed for seasonal adjustment.

The seasonal-irregular component $(si)_t$ (table B 8) is determined in the second step as the ratio of the calendar-adjusted time series and the estimated trend-cycle component. This seasonal-irregular component is smoothed using a monthly or quarterly-specific moving average in the third step. After normalisation, the outcome is an estimation of the seasonal component s_t (table B 10).

For the seasonal component, three-period simple moving averages of a simple moving average of odd length are used as smoothing filters. This means that this smoothing filter is based on a combination of two simple moving averages. The 3×9-moving seasonal filter, for example, has the filter weights

$$\left\{\frac{1}{27}, \frac{2}{27}, \frac{3}{27}, \frac{3}{27}, \frac{3}{27}, \frac{3}{27}, \frac{3}{27}, \frac{3}{27}, \frac{3}{27}, \frac{2}{27}, \frac{1}{27}\right\},$$

which originate from the systematic structure shown in Figure 3.9. This smoothing filter has a validity window of eleven years since, contrary to the Henderson trend filter (see below), the moving seasonal filter is calculated on a monthly or quarterly-specific basis.

The symmetry of the smoothing filter prevents a phase shift, which would cause the seasonally adjusted time series to have different turning points from that of the unadjusted time series.

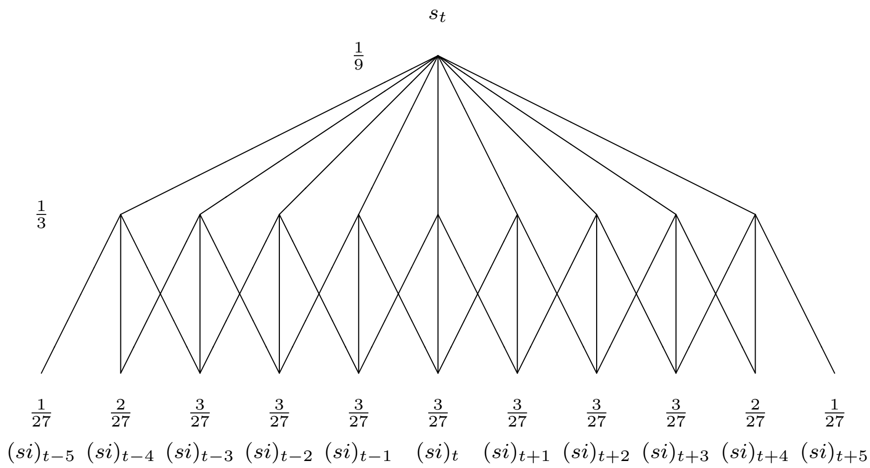


Figure 3.9: Smoothing Filter Weights of the 3×9-Moving Seasonal Filter

Using this estimate of the seasonal component, the seasonally and calendar-adjusted time series (table B 11) is determined in step four. It is noteworthy that smoothness of the seasonally adjusted time series $a_t = u_t/s_t$ from Equation (3.2), is no criterion for its quality. As it can be seen from the same equation the seasonally adjusted time series $a_t = c_t \cdot i_t$ contains the irregular component per definitionem. Also it should be noted that the movements of its period-to-period changes $\Delta_t = a_t/a_{t-1} - 1$ from Equation (3.3) are at least partly random noise.

The irregular component i_t (table B 13) comprises all random and exceptional fluctuations as well as the errors of the other components. It is derived as the ratio of the seasonally and calendar-adjusted time series and the trend-cycle component. If extreme values are recorded for the irregular component, these seasonal-irregular components are given a lower weighting (table B 17) and are replaced by adjacent values – as exemplarily shown in Figure 3.10. Here, full weighting is used within 1.5 times the standard deviation of the expectation ($\mathrm{E}(i_t) = 1$) and no weighting is used beyond 2.5 times the standard deviation.

Based on this provisional weighting of extreme values in section B, the whole process of smoothing out the trend and seasonal component from the calendar and outlier-adjusted time series is repeated in section C in order to determine the final weighting of extreme values (table C 17). Lastly, the final estimate of the components (tables D 10, D 11, D 12 and D 13) is calculated in section D using the irregular-weighted time series whose weights were calculated in section C.

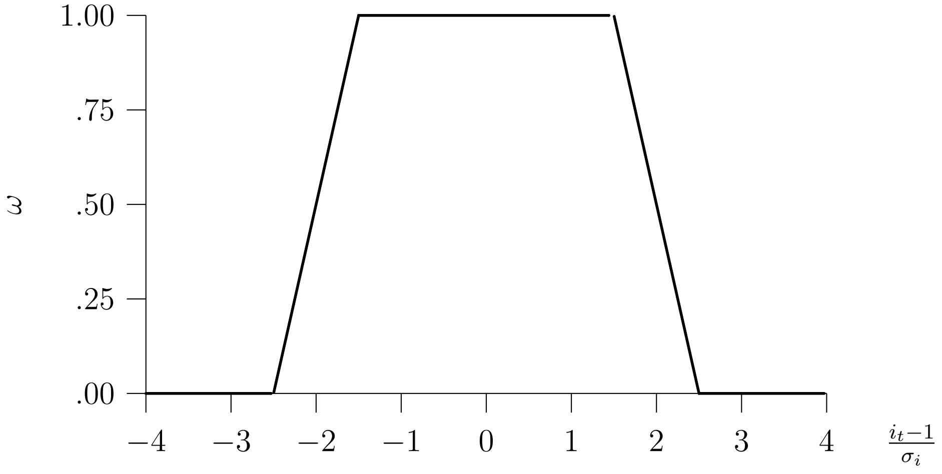


Figure 3.10: Down-Weighting of Extreme Seasonal Factors

For the trend-cycle component (table D 12), Henderson smoothing filters of odd length are applied to the seasonally and calendar-adjusted time series. The filter weights $\omega_{t+\tau}$ solve the minimisation problem $\sum_{-n}^{n}(\Delta^3\omega_{t+\tau})^2$ subject to $\sum_{-n}^{n}\omega_{t+\tau} = 1$, $\sum_{-n}^{n} t \cdot \omega_{t+\tau} = 0$ and $\sum_{-n}^{n} t^2 \cdot \omega_{t+\tau} = 0$. Analytically, for period $t+\tau$ ($\tau \in [-n, n]$) the following formula results, where $m = n+2$:

$$\omega_{t+\tau} = \frac{315[(m-1)^2-\tau^2][m^2-\tau^2][(m+1)^2-\tau^2][(3m^2-16)-11\tau^2]}{8m(m^2-1)(4m^2-1)(4m^2-9)(4m^2-25)}.$$

The 13-term Henderson trend filter, for instance, has the filter weights

$$\left\{-\frac{325}{16.796}, -\frac{468}{16.796}, \frac{0}{16.796}, \frac{1.100}{16.796}, \frac{2.475}{16.796}, \frac{3.600}{16.796}, \frac{4.032}{16.796}, \right.$$

$$\left. \frac{3.600}{16.796}, \frac{2.475}{16.796}, \frac{1.100}{16.796}, \frac{0}{16.796}, -\frac{468}{16.796}, -\frac{325}{16.796}\right\}.$$

This smoothing filter extends over $2n+1 = 13$ successive periods ($n = 6$).

Figure 3.11 shows the smoothing filter properties in the frequency domain of a combined monthly 3×9-moving seasonal filter and 13-term Henderson trend filter using the squared gain. This smoothing filter eliminates the monthly frequencies while the spectral mass is heightened around the monthly frequencies.

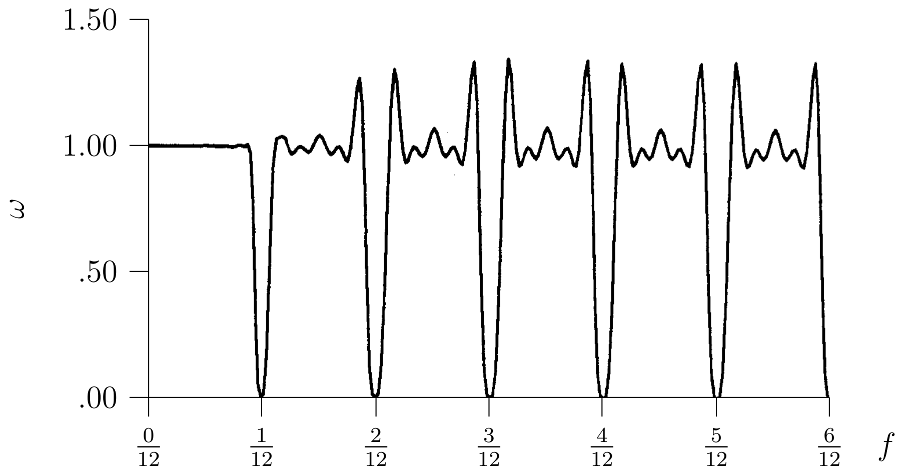


Figure 3.11: Squared Gain of the Combined Smoothing Filter

Appendix B: Real Time Data Properties

Data Description

Gross domestic product in constant prices in million DM was calculated with base year 1991 up to the first publication of reporting date first quarter 1999. From then onwards it was based on the year 1995 and from reporting date fourth quarter 2001 onwards it is given in million Euro. The base year 2000, first introduced for reporting date first quarter 2005, marked the changeover to previous year's price basis chain-linked Laspeyres volume indices with the annual overlap technique. To account for this, all figures are converted to Euro if applicable, and fixed price basis volumes are re-indexed with base year 2000 = 100 to match the current chain index. Working-day adjustment is performed beforehand, utilising monthly indicator time series.

Figure 3.12 shows the real time data vintages of gross domestic product on the logarithmic scale for the last two years. The vintage graph exhibits the effect of revisions to the time series. One can see for example that the growth path (on the linear scale) has been revised continuously upwards in 2006 since new data became available.

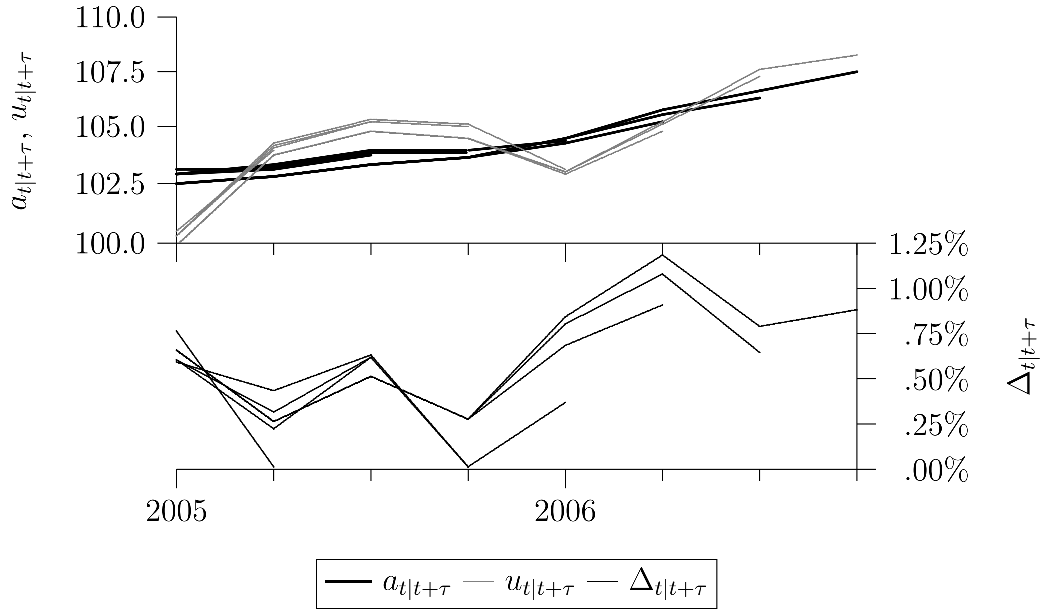


Figure 3.12: Real Time Data Vintages of Real Gross Domestic Product

Although revisions of real time data decline over time, they do not vanish as is shown in Figure 3.13 which plots average absolute revisions of real time data of gross domestic product against τ. The decline is to be expected given the increase of available data with the passing of time. That revisions do not vanish completely is due to benchmark revisions, e.g. changes in classifications.

The most recent vintage of the seasonally adjusted and unadjusted time series as well as the seasonal component is drawn in Figure 3.14. The axis of the seasonally adjusted and unadjusted time series is logarithmically scaled and the one of the seasonal component is linearly scaled.

Employment was measured in 1,000 persons until the end of 1998 without persons in marginal employment, and with these afterwards. While prior to reporting month June 2000 the microcensus of the Federal Statistical Office was used as one important source of the base data, since then it has been replaced by the employment statistics of the Federal Employment Agency. These effects are captured by linking the vintages to show no revisions, i.e. monthly specific chain factors are calculated that neutralise both these effects in the respective months. Time series are to be found in Figure 3.15. Note that both axes are linearly scaled because employment is not an index unlike the other indicators.

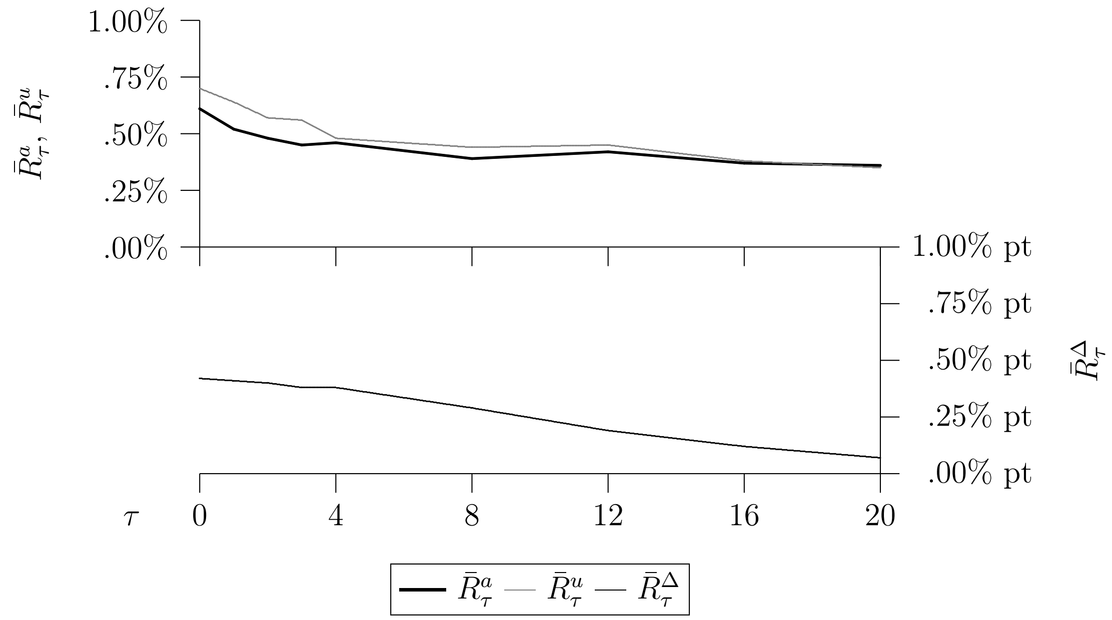


Figure 3.13: Revisions of Real Time Data of Real Gross Domestic Product

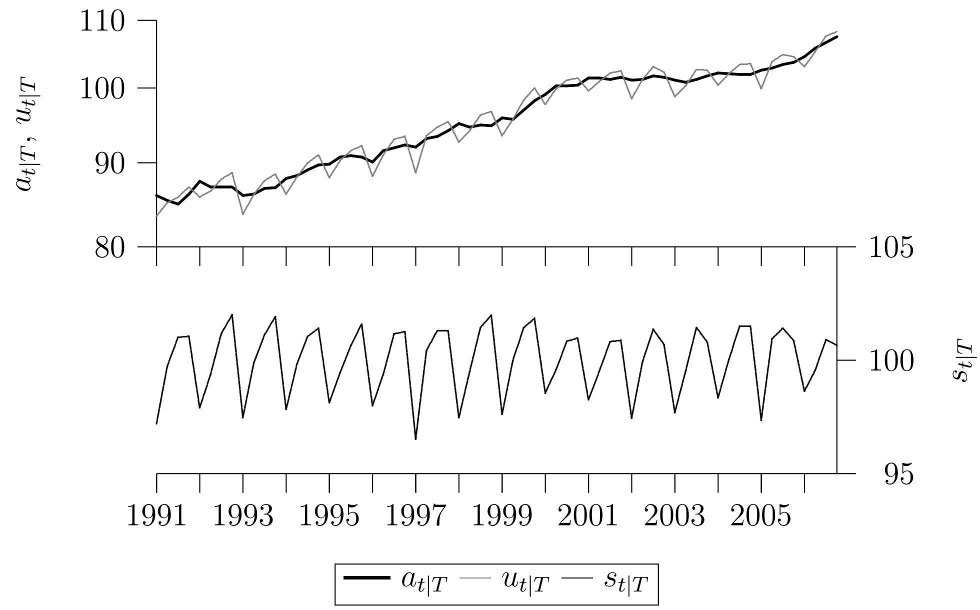


Figure 3.14: Time Series of Real Gross Domestic Product

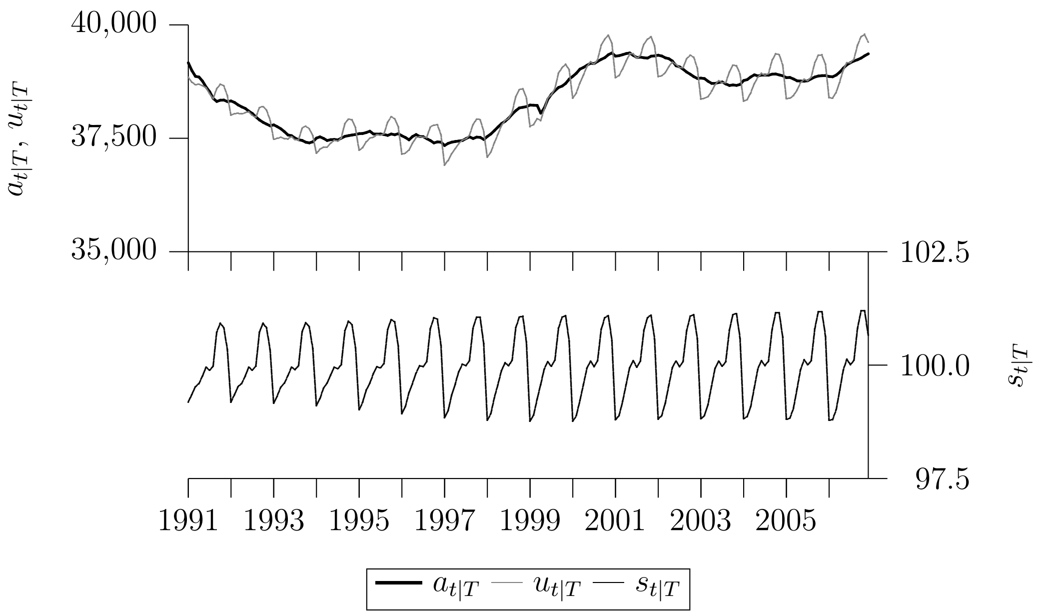


Figure 3.15: Time Series of Employment

Output in and orders received by the manufacturing sector both begin as value indices with base year 1991 = 100 up to reporting month May 1998 when the base year was changed to 1995 = 100. The current base year 2000 = 100 was introduced for output in reporting month December 2003 and for orders received in January 2003. All figures used for this analysis are based on the year 2000 = 100. Figures 3.16 and 3.17 show the time series.

Retail trade turnover is evaluated as an index in current prices excluding value added tax. Its base year was changed in reporting month January 1997 from 1994 = 100 to 1995 = 100, in August 2002 to 2000 = 100 and in April 2005 to 2003 = 100. For this analysis the last base year, 2003 = 100, is the reference, too. Furthermore, the very first estimate based only on six out of sixteen federal states is used throughout. Time series are given in Figure 3.18.

Table 3.6 collects the user settings of seasonal adjustment. In particular, the chosen ARIMA model, the seasonal and trend filter lengths and the sigma limits (cf. Appendix A for the meaning of each of these settings).

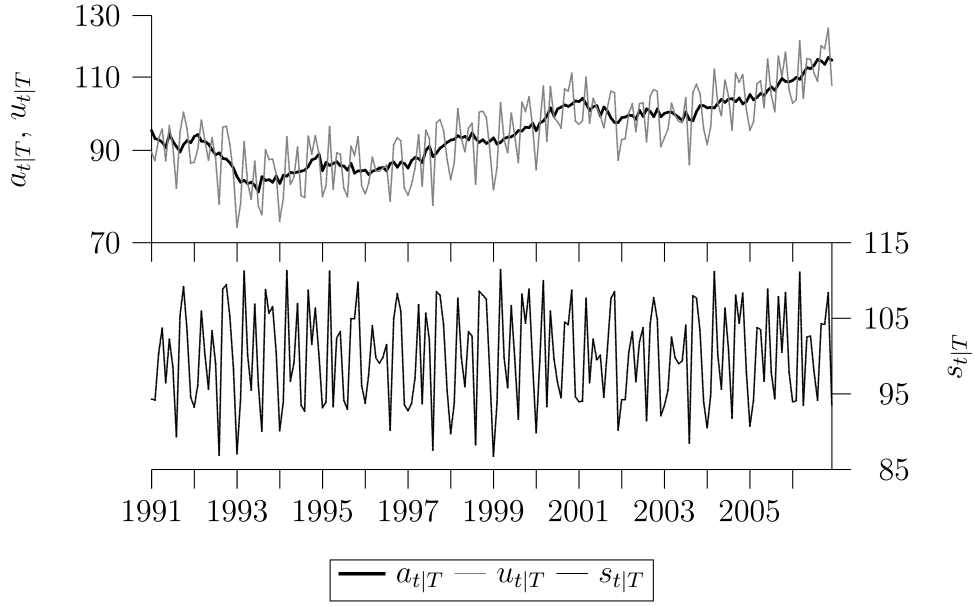


Figure 3.16: Time Series of Output in the Manufacturing Sector

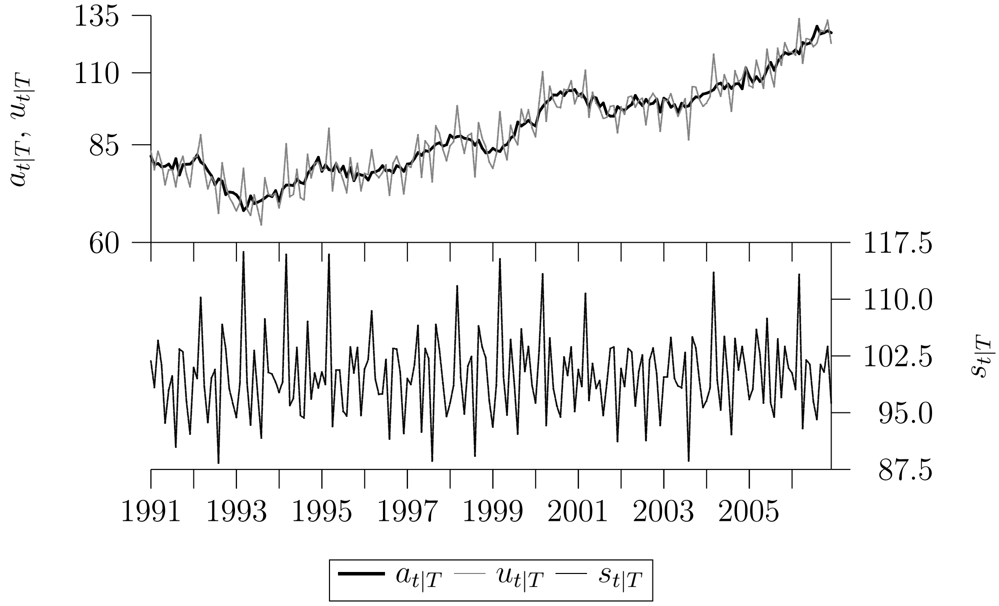


Figure 3.17: Time Series of Orders Received by the Manufacturing Sector

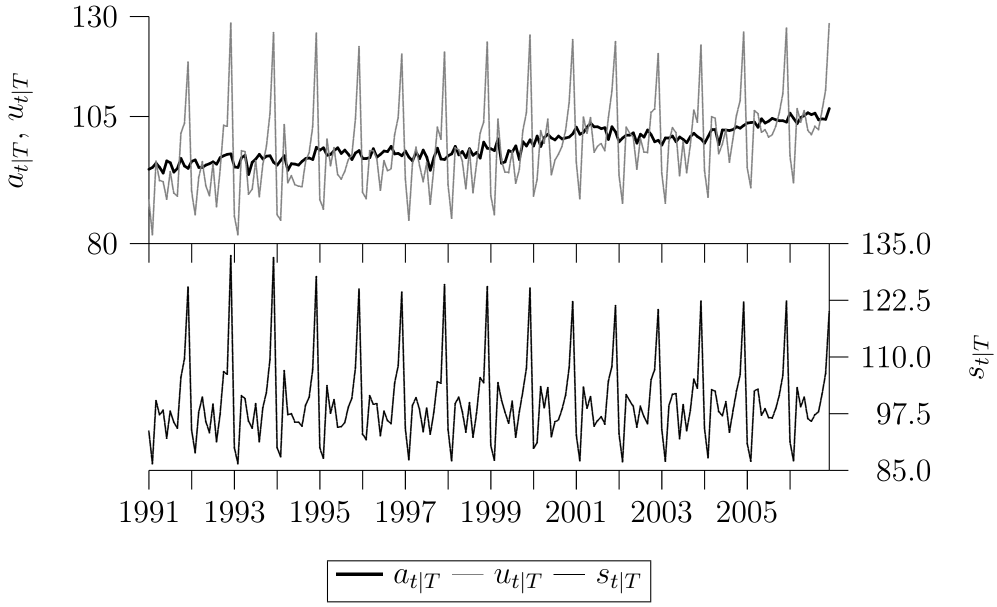


Figure 3.18: Time Series of Retail Trade Turnover

Table 3.6: User Settings of Seasonal Adjustment of Time Series†

Time Series	ARIMA Model	Filter Lengths: Seasonal	Filter Lengths: Trend	Sigma Limits
Gross Domestic Product	(0,1,0)(0,1,1)	3×9	5	2.0 / 3.0
Employment	(0,2,2)(0,1,1)	3×5	13	2.5 / 3.5
Output	(0,1,1)(0,1,1)	3×9	17	2.0 / 3.0
Orders Received	(0,1,1)(0,1,1)	3×15	13	1.5 / 2.5
Retail Trade Turnover	(0,1,1)(0,1,1)	3×9	17	2.3 / 3.0

† Seasonal filters are applied separately to each month and quarter, respectively, and may therefore differ from one another. In these cases the most frequently used length is stated in the table.

Data Tables

A breakdown of numeric values of time series specific average absolute percent ex ante-forecast errors over the last three years of Figure 3.3 is given in Table 3.7.

Table 3.7: Average Absolute Percent Ex Ante-Forecast Errors

Time Series	Last 3 Years	Last Year	Last−1 Year	Last−2 Year
Gross Domestic Product	1.12%	1.91%	1.06%	0.40%
Employment	0.50%	0.33%	0.55%	0.62%
Output	2.36%	2.74%	3.06%	1.27%
Orders Received	2.48%	2.65%	3.13%	1.65%
Retail Trade Turnover	1.37%	0.89%	1.50%	1.72%

Numeric values of time series specific average absolute revisions of seasonally adjusted real time data, seasonal adjustment and unadjusted real time data and their corresponding period-to-period changes, drawn in Figure 3.4, are given in Table 3.8.

Table 3.8: Average Absolute Revisions

	Levels			Period-to-period changes		
Time Series	$\bar{R}_i^a$	$\tilde{R}_i^s$	$\bar{R}_i^u$	$\bar{R}_i^a$	$\tilde{R}_i^s$	$\bar{R}_i^u$
Gross Domestic Product	0.60%	0.18%	0.70%	0.40% pt	0.28% pt	0.70% pt
Employment	1.24%	0.06%	1.25%	0.07% pt	0.03% pt	0.13% pt
Output	1.23%	0.44%	1.26%	0.94% pt	0.48% pt	0.92% pt
Orders Received	1.08%	0.60%	0.75%	1.25% pt	0.83% pt	0.85% pt
Retail Trade Turnover	0.84%	0.43%	0.72%	0.70% pt	0.43% pt	0.64% pt

In Table 3.9 numeric values are collected for the standard deviations of the irregular and the seasonal component from Figure 3.5 along with average duration of run and periods of cyclical dominance from Figure 3.6.

Table 3.9: Time Series Properties and X-11 Statistics

Time Series	σ_i	σ_s	ADR	PCD
Gross Domestic Product	0.18%	1.45%	2.42	1
Employment	0.08%	0.72%	3.44	1
Output	0.82%	5.52%	1.68	3
Orders Received	1.35%	4.21%	1.74	3
Retail Trade Turnover	0.79%	8.63%	1.72	7

Appendix C: Estimation of Elasticities

Pseudocode

For fixing the model and keeping outliers throughout revision analysis, all parameters need to be `f`-ixed in the specification file using both the `b`, and `ar` and `ma`-tags in the `regression` and `arima` section, respectively. Additionally, the `outlier` section needs to be deleted as all outliers should be included in the `regression` section. The pseudocode for the calculation of seasonally adjusted figures, and their revisions and elasticities reads as follows.

Pseudocode	Input	Output	Reference
Seasonally adjust real time data	$u_{t\|t}$	$a_{t\|t}$	Eq. (3.2)
Seasonally adjust historic data	$u^t_{t\|T}$	$a^t_{t\|T}$	Tab. 3.2
Calculate period-to-period changes	$a_{t\|t}$	$\Delta_{t\|t}$	Eq. (3.3)
	$a^t_{t\|T}$	$\Delta^t_{t\|T}$	
Calculate revisions	$a_{t\|t}$, $a_{t\|T}$	r^a_t	Eq. (3.4)
	$a^t_{t\|T}$, $a_{t\|T}$	$\tilde{r}^s_t$	
	$u_{t\|t}$, $u_{t\|T}$	r^u_t	
	$\Delta_{t\|t}$, $\Delta_{t\|T}$	r^Δ_t	Eq. (3.6)
Estimate marginal effects	r^a_t, $\tilde{r}^s_t$, r^u_t	β^s, β^u	Eq. (3.8)
Calculate elasticities	β^s, β^u, $\bar{R}^a$, $\bar{\tilde{R}}^s$, $\bar{R}^u$	ε^s, ε^u	Eq. (3.9)

Delta Method

Standard errors of elasticities and their ratio are calculated using the Delta Method. The Delta Method is based on a first-order Taylor approximation of the non-linear expression around its expectation. Let $\mathbf{X}$ be the vector of known random variables and $g(\mathbf{X})$ be the desired transformation into a new random variable Y. It then follows that $Y = g(\mathbf{X}) \approx g(\mathrm{E}(\mathbf{X})) + \mathbf{J} \cdot (\mathbf{X} - \mathrm{E}(\mathbf{X}))$, where $\mathbf{J} = \partial g(\mathbf{X})/\partial \mathbf{X}$ is the Jacobian matrix. The expectation and variance of Y are $\mathrm{E}(Y) \approx g(\mathrm{E}(\mathbf{X}))$ and $\mathrm{Var}(Y) \approx \mathbf{J} \cdot \mathrm{Var}(\mathbf{X}) \cdot \mathbf{J}'$, respectively. For the calculation of variances of elasticities and their ratio, estimated slope coefficients and average absolute revisions were used. Under the assumption of $\mathrm{Cov}(\boldsymbol{\beta}_\mathbf{i}, \bar{\mathbf{R}}_\mathbf{i}) = \mathbf{0}$, $\mathrm{Var}(\mathbf{X})$ can be block-diagonal partitioned and the variance of Y can be written as $\mathrm{Var}(Y) \approx \mathbf{J}_{\boldsymbol{\beta}} \cdot \mathrm{Var}(\boldsymbol{\beta}_\mathbf{i}) \cdot \mathbf{J}'_{\boldsymbol{\beta}} + \mathbf{J}_{\bar{\mathbf{R}}} \cdot \mathrm{Var}(\bar{\mathbf{R}}_\mathbf{i}) \cdot \mathbf{J}'_{\bar{\mathbf{R}}}$, where $\mathrm{Var}(\bar{\mathbf{R}}_\mathbf{i}) = 1/T \cdot \mathrm{Var}(\mathbf{R}_\mathbf{i})$.

In the case of the variance of the elasticity of seasonal adjustment revisions, $\mathbf{X}$ reads

$$\mathbf{X} = \begin{bmatrix} \beta_i^s & \bar{R}_i^a & \bar{\bar{R}}_i^s \end{bmatrix}'.$$

The first partial derivatives of $g(\mathbf{X})$ as in Equation (3.9) with respect to $\mathbf{X}$ are

$$\mathbf{J} = \begin{bmatrix} \frac{\bar{\bar{R}}_i^s}{\bar{R}_i^a} & -\beta_i^s \cdot \frac{\bar{\bar{R}}_i^s}{\left(\bar{R}_i^a\right)^2} & \beta_i^s \cdot \frac{1}{\bar{R}_i^a} \end{bmatrix}.$$

Analogously, replace β_i^s and $\bar{\bar{R}}_i^s$ with β_i^u and $\bar{R}_i^u$, respectively, for the elasticity of unadjusted real time data revisions.

In the same manner one finds $\mathbf{X}$ and $\mathbf{J}$ for the ratio of these elasticities.

$$\mathbf{X} = \begin{bmatrix} \beta_i^s & \beta_i^u & \bar{\bar{R}}_i^s & \bar{R}_i^u \end{bmatrix}'$$

$$\mathbf{J} = \begin{bmatrix} -\frac{\beta_i^u \cdot \bar{R}_i^u}{\left(\beta_i^s\right)^2 \cdot \bar{\bar{R}}_i^s} & \frac{\bar{R}_i^u}{\beta_i^s \cdot \bar{\bar{R}}_i^s} & -\frac{\beta_i^u \cdot \bar{R}_i^u}{\beta_i^s \cdot \left(\bar{\bar{R}}_i^s\right)^2} & \frac{\beta_i^u}{\beta_i^s \cdot \bar{\bar{R}}_i^s} \end{bmatrix}$$

Chapter 4

A Solution to the Problem of Too Many Instruments in Dynamic Panel Data GMM

4.1 The Problem of Too Many Instruments

Dynamic panel data (DPD) models have become increasingly popular in the last two decades. Nowadays the availability of micro level data, such as of firms or banks, enables researchers to identify economic relationships at a disaggregate level. Hence, the serious problem of aggregation bias (Lippi and Forni, 1990) can be avoided. However, the solution is not without a drawback: DPD bias. As Nickel (1981) has shown, the Least Squares Dummy Variables (LSDV) estimator has a non-vanishing bias for small T and large N. Anderson and Hsiao (1982) were the first to propose an unbiased DPD estimator with the notable trade-off between lag depth and sample size. It was not until Holtz-Eatkin et al. (1988) that an unbiased DPD estimator was constructed based on Generalised Method of Moments (GMM) (Hansen, 1982). The breakthrough came with Difference GMM by Arellano and Bond (1991), and System GMM by Arellano and Bover (1995) and Blundell and Bond (1998). In the meantime, Kiviet (1995) proposed a corrected LSDV estimator for balanced panels. However, one issue with regard to DPD GMM still remains unresolved; the number of instruments grows quadratically in T and GMM becomes inconsistent as the number of instruments diverges, thus begging the question "what is the optimal set of instruments?"

Roodman (2009) addresses the problem of too many instruments. Increasing the sample size causes the number of instruments to proliferate as DPD GMM generates one instrument for each time period and lag available. Currently, there are two techniques in use to reduce the instrument count. One of them is limiting the lag depth, the other one is "collapsing" the instrument set. The former implies a selection of certain lags to be included in the instrument set, making the instrument count linear in T. The latter embodies a different belief about the orthogonality condition: it no longer needs to be valid for any one time period but still for each lag, again making the instrument count linear in T. A combination of both techniques makes the instrument count invariant to T. These transformations are deterministic ones of the instrument matrix, i.e. the transformation matrix consists of zeroes and ones. Besides the fact that no widely accepted rule of thumb for the instrument count exists, by choosing one of the aforementioned approaches, the researcher decides which transformation is to be used for the data. The point in question is, "can we let the data decide how the transformation matrix should look?" The answer to this question is found by means of factor analysis of the instrument set and is shown to be "yes, we can." The resulting DPD GMM estimator is characterised by both a lower bias and a lower root mean squared error (RMSE) than the standard techniques.

The remainder of the chapter is organised as follows. Section 4.2 introduces the new estimation technique based on factorised instruments. Monte Carlo results for this estimator are presented in Section 4.3. The final section concludes.

4.2 A Solution to this Problem

Consider an autoregressive panel model of order one for the endogenous variable $y_{i,t}$, where $\alpha_i = \alpha + \eta_i$ is a fixed effect and $\varepsilon_{i,t}$ is the error term.

$$y_{i,t} = \alpha + \beta y_{i,t-1} + \eta_i + \varepsilon_{i,t} \tag{4.1}$$

The autoregressive parameter β of Equation (4.1) is estimated with DPD GMM in first differences ($\Delta y_{i,t} = \beta \Delta y_{i,t-1} + \Delta\varepsilon_{i,t}$). This will be treated here exclusively without loss of generality but for simplicity of exposition. The standard instrument set $\mathbf{Z}$ consists of lagged values of the endogenous variable, which are uncorrelated with the first differences of the error term.

$$\mathrm{E}(\mathbf{Z}'\boldsymbol{\Delta\varepsilon}) = \mathbf{0} \tag{4.2}$$

First, the conditions for consistency of the aforementioned techniques, along with a whole class of transformations, to reduce the instrument count are verified in the following theorem. Unlike other authors, who derive the limited or collapsed instrument set from first principles by considering interpretable orthogonality conditions, this chapter applies transformation matrices to the standard instrument set which yield the desired results (cf. Appendix B). Proofs for this and the following theorem are to be found in Appendix A.

Theorem 4.1. *Let Equation (4.2) be valid. Then* $\mathrm{E}(\mathbf{Z}^{*\prime}\boldsymbol{\Delta\varepsilon}) = \mathbf{0}$ *with* $\mathbf{Z}^* = \mathbf{ZF}$ *for any deterministic transformation matrix* $\mathbf{F}$.

It follows from Theorem 4.1 that limiting the lag depth, collapsing the instrument set or both are valid transformations for consistent estimation of the parameter of interest. Moreover, any transformation, no matter if it lacks a sensible interpretation, satisfies the conditions of the theorem as long as it is deterministic.

Second, the aim of this chapter is to introduce a new technique rather than to evaluate standards already in use. Hence, the focus here lies on stochastic transformations instead of deterministic ones. In order to solve the problem of instrument proliferation, this chapter suggests the application of factor analysis – more precisely for the case in hand – principal components analysis (PCA) to the instrument set. PCA extracts the largest eigenvalues of the estimated covariance matrix of $\mathbf{Z}$ and assembles the corresponding eigenvectors in the matrix of component loadings $\mathbf{F}^*$, the transformation matrix. In this case, the transformation matrix is stochastic and Theorem 4.1 is no longer applicable. However, Theorem 4.2 provides a solution.

Theorem 4.2. *Let $y_{i,t-1-\ell}$, $\ell = 1, 2, \ldots$ (the elements of the $\mathbf{Z}$ matrix) and $\Delta\varepsilon_{i,t}$ be independent random variables for all i and t. Then $\mathrm{E}(\mathbf{Z}^{**\prime}\boldsymbol{\Delta\varepsilon}) = \mathbf{0}$ with $\mathbf{Z}^{**} = \mathbf{Z}\mathbf{F}^*$, where $\mathbf{F}^*$ is the matrix of component loadings from PCA of $\widehat{\mathrm{Var}}(\mathbf{Z})$.*

Theorem 4.2 is both more general and more specific than Theorem 4.1. The fact that it also holds true for deterministic $\mathbf{F}^* = \mathbf{F}$ makes it more general. It is more specific in the sense that it requires independence of $y_{i,t-1-\ell}$ and $\Delta\varepsilon_{i,t}$ which is a stronger property than uncorrelatedness. This assumption is not too strong if the error term is thought of as being an exogenous shock.

4.3 Performance of Factorised Instruments

Judson and Owen (1999) provide Monte Carlo evidence that GMM is superior to other estimation techniques when it comes to DPD. Among others, their findings are: OLS produces biased estimates even for large T, the bias of LSDV decreases with T but may still be up to 20% of the true value even when $T = 30$, and also that the LSDV bias increases with the true value of the autoregressive parameter. Additionally, OLS is upward biased while LSDV is downward biased. Windmeijer (2005) adds to this list that GMM becomes more efficient when the lag depth is limited, and thus fewer instruments are employed in the estimation.

Table 4.1 and Figures 4.1, 4.2 and 4.3 present biases, RMSEs and standard deviations (SDs) from a Monte Carlo simulation of a one-step estimation of Equation (4.1) with parameter values of β in the range from close to zero to close to one. $\varepsilon_{i,t}$ is assumed to be standard normal, as is α_i. N is fixed at 100, T is 10, 20 and 30, respectively (large N, small T). The pre-sample period length is 30. The standard instrument set is either taken as it is, limited, collapsed or both, and additionally PCA has been applied to all four variants. The experiment is repeated 1,000 times. The derivation of the GMM estimator and the test statistics can be found in Appendix C.

The results confirm the findings of Judson and Owen (1999) and Windmeijer (2005). In addition, factorised instruments outperform all other techniques by having both a lower bias and RMSE, however, there are a few exceptions when $T = 10$. In general, factorisation of the limited and collapsed instrument set results in the lowest bias, while factorisation of the collapsed but unlimited instrument set yields the lowest RMSE. Biases are zero to the second decimal place or in relative terms less than 1%, RMSEs are zero to the first decimal place. SDs reveal the frequently cited result that GMM is potentially more volatile than least squares estimators. This is even more the case if factorised instruments are used. The well-known trade-off between a loss of efficiency due to the neglect of theoretically valid instruments and a reduction of bias by avoiding instrument proliferation arises.

The advantage of factorised instruments over standard ones is the condensation of the informational content of the instrument set into a much lower number of instruments employed in the estimation thus lowering the risk of overfitting endogenous variables but retaining almost all information. The next best approach is standard GMM with the instrument set being both limited and collapsed. Acceptable results can also be derived from a collapsed but unlimited instrument set in standard GMM. Limiting the lag depth on the one hand is a good idea as even if the autoregressive parameter is high, serial correlation will be low after a few periods and deeper lags are weak instruments, adding almost no new information for estimation. Collapsing the instrument set on the other hand also condenses the information in the instrument set into a lower number of instruments. The techniques most frequently used in applied DPD research, the untransformed in-

Table 4.1: Bias, RMSE and SD for $\beta = .2$ and $\beta = .8$[†]

		$T = 10$		$T = 20$		$T = 30$	
Method	Statistic	$\beta = .2$	$\beta = .8$	$\beta = .2$	$\beta = .8$	$\beta = .2$	$\beta = .8$
Least Squares							
OLS	Bias	$+.477$	$+.180$	$+.477$	$+.180$	$+.477$	$+.180$
	RMSE	.478	.180	.478	.180	.478	.180
	SD	.035	.005	.032	.004	.031	.003
LSDV	Bias	$-.136$	$-.243$	$-.064$	$-.111$	$-.042$	$-.070$
	RMSE	.140	.245	.068	.113	.045	.071
	SD	.033	.031	.022	.019	.018	.014
Standard GMM							
Untransformed	Bias	$-.080$	$-.539$	$-.146$	$-.624$	$-.199$	$-.681$
	RMSE	.101	.555	.151	.628	.201	.683
	SD	.062	.133	.038	.066	.029	.044
Limited (Ltd.)	Bias	$-.061$	$-.506$	$-.114$	$-.580$	$-.157$	$-.633$
	RMSE	.089	.528	.121	.585	.160	.635
	SD	.065	.151	.039	.075	.031	.051
Collapsed (Col.)	Bias	$-.014$	$-.373$	$-.017$	$-.296$	$-.017$	$-.257$
	RMSE	.070	.435	.047	.325	.039	.275
	SD	.069	.225	.043	.133	.035	.098
Ltd. & Col.	Bias	$-.001$	$-.172$	$-.007$	$-.159$	$-.007$	$-.137$
	RMSE	.071	.297	.044	.205	.036	.166
	SD	.071	.243	.044	.129	.035	.094
Factorised GMM							
Untransformed	Bias	$-.325$	$-.706$	$-.463$	$-.826$	$-.502$	$-.856$
	RMSE	.550	.913	.632	.945	.607	.949
	SD	.444	.579	.430	.459	.341	.408
Limited (Ltd.)	Bias	$-.165$	$-.534$	$-.300$	$-.646$	$-.399$	$-.760$
	RMSE	.305	.769	.447	.781	.501	.861
	SD	.256	.553	.331	.439	.303	.405
Collapsed (Col.)	Bias	$+.004$	$-.026$	$+.003$	$-.007$	$+.004$	.000
	RMSE	.059	.189	.035	.077	.029	.048
	SD	.059	.188	.035	.077	.028	.048
Ltd. & Col.	Bias	$+.002$	$+.005$	$+.002$	$-.002$	$+.003$	.000
	RMSE	.067	.217	.037	.084	.031	.055
	SD	.067	.217	.037	.084	.030	.055

[†] For the sake of brevity, results for values of the autoregressive parameter other than $\beta = .2$ and $\beta = .8$ are not displayed here. The results obtained for these values are similar to those presented above.

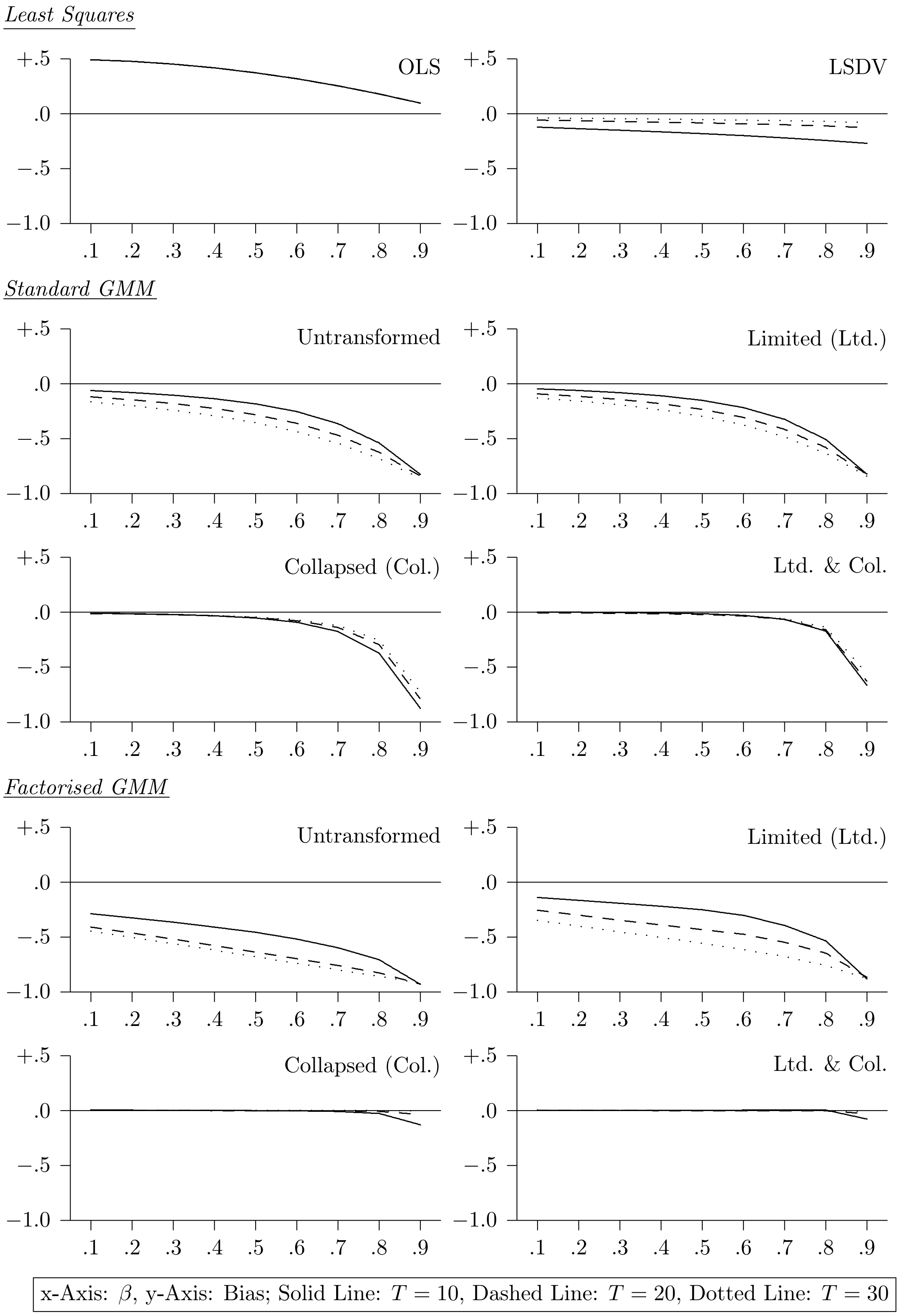


Figure 4.1: Biases from a Monte Carlo Simulation

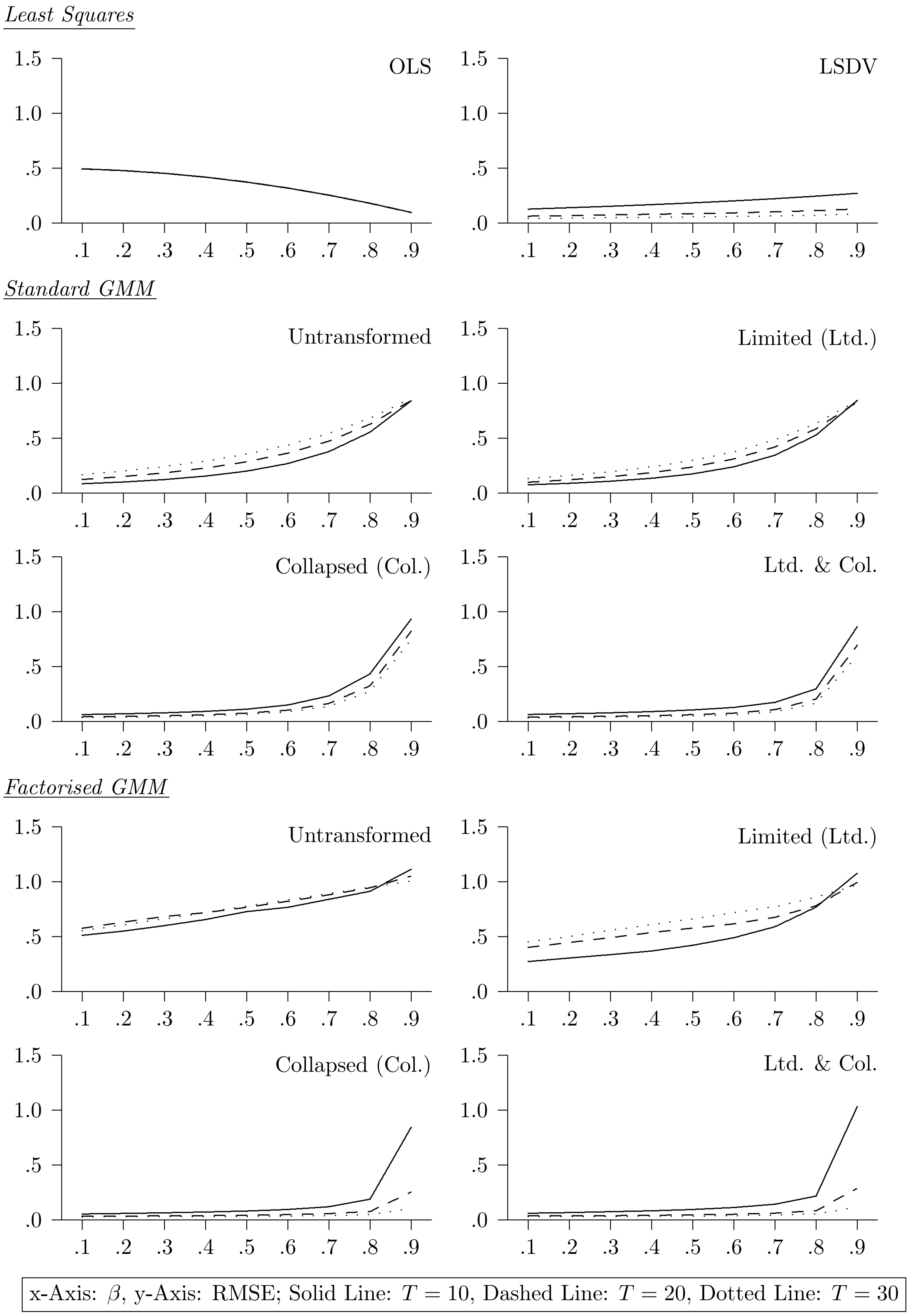
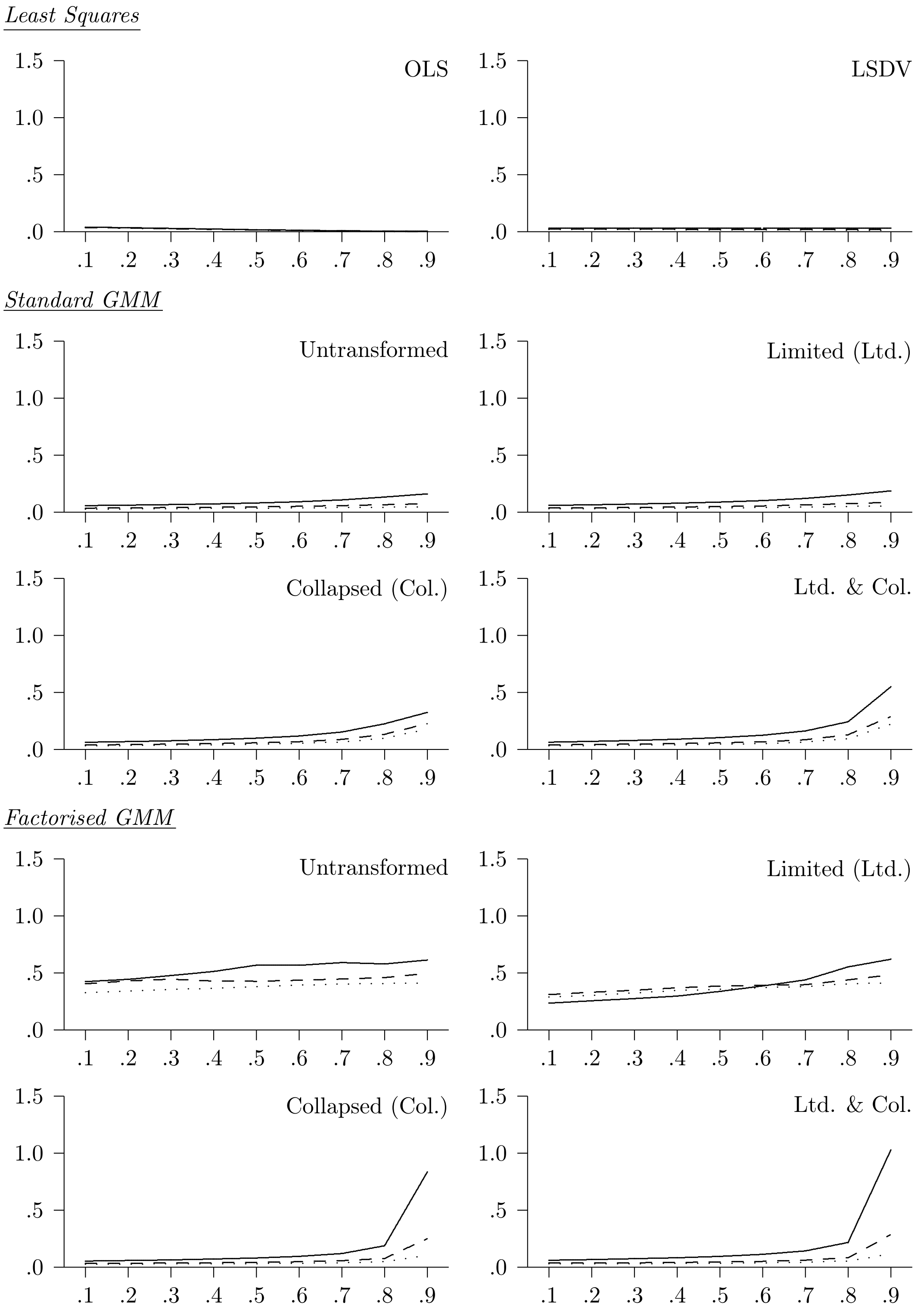


Figure 4.2: RMSEs from a Monte Carlo Simulation



x-Axis: β, y-Axis: SD; Solid Line: $T = 10$, Dashed Line: $T = 20$, Dotted Line: $T = 30$

Figure 4.3: SDs from a Monte Carlo Simulation

strument set and the limited one in standard GMM, are the worst choices, that is apart from the factorised variants of them. Both techniques are significantly downward biased (which becomes even worse, the higher T is), although the estimate still has the correct sign. Performance of their factorised variants is unacceptable; not even the correct sign can be expected.

Explanations for the failure of the standard techniques can be found with recourse to the Sargan (1958) test of overidentifying restrictions (cf. Table 4.2 and Figure 4.4). The failure of the factorised variants can be traced back to PCA and the Kaiser-Meyer-Olkin (Kaiser, 1970) measure of sampling adequacy (MSA) (cf. Table 4.3 and Figure 4.5). Through testing for weak instruments according to Staiger and Stock (1997), more evidence is found why both the standard techniques and the factorised variants do not perform particularly well (cf. Table 4.4 and Figure 4.6).

Table 4.2 shows the number of instruments employed in the estimation for each of the methods used and the proportions for which the validity of the overidentifying restrictions have been rejected at the nominal 5% significance level. It should be borne in mind that the power of the test is not weakened by many instruments. For limited instrument sets, the number of lags employed is set to be half of the available lags; for factorised instrument sets, the number of retained components has been fixed. Both choices are to a certain extent arbitrary.

Table 4.2: Instrument Count J and Rejection Frequency of Valid Instruments

	$T = 10$			$T = 20$			$T = 30$		
Method	J	$\beta = .2$	$\beta = .8$	J	$\beta = .2$	$\beta = .8$	J	$\beta = .2$	$\beta = .8$
Standard GMM									
Untransformed	36	.103	.202	171	.176	.400	406	.318	.605
Limited (Ltd.)	26	.096	.181	126	.140	.365	301	.228	.568
Collapsed (Col.)	8	.091	.166	18	.077	.169	28	.092	.185
Ltd. & Col.	4	.047	.097	9	.069	.096	14	.074	.099
Factorised GMM									
Untransformed	3	.080	.076	4	.064	.057	5	.070	.064
Limited (Ltd.)	3	.100	.109	4	.063	.064	5	.076	.072
Collapsed (Col.)	2	.000	.000	3	.000	.000	4	.000	.000
Ltd. & Col.	2	.000	.001	3	.000	.000	4	.000	.000

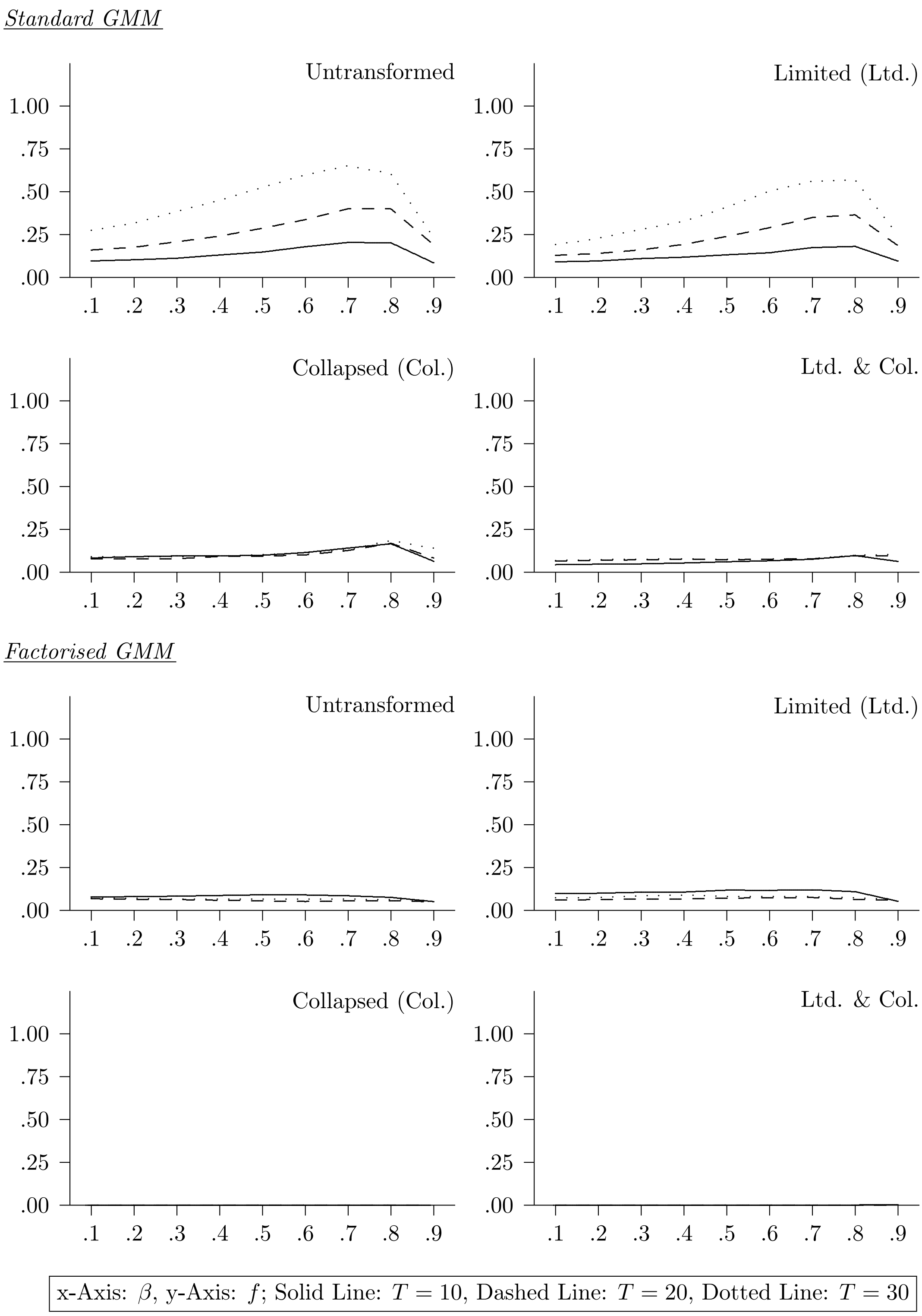


Figure 4.4: Rejection Frequency f of Valid Instruments

Standard GMM with the untransformed or limited instrument set generates invalid overidentifying restrictions in an unacceptably high number of cases. This is due to the impossibility of fulfilling all restrictions simultaneously owing to the large number of instruments and the resulting overfitting of endogenous variables. Probabilities of rejection increase with β as well as with T (cf. Figure 4.4). As it is known a priori that the null hypothesis of valid instruments or overidentifying restrictions is true in all cases, severe size distortions of the test become visible. While the test of the factorised variants of the collapsed (and limited) instrument set is undersized, rejecting the null hypothesis in virtually none of the cases, all tests of other instrument sets are oversized, some rather heavily.

Table 4.3 reports the explained variance and MSA from PCA. The explained variance states the proportion of the instrument set's variance that can be explained by the retained components. MSA is a statistical criterion to judge the adequacy of the covariance matrix to be factorised; the closer it gets to one, the better. A value in the .90s is regarded as being "marvellous" in the literature (Kaiser and Rice, 1974).

Table 4.3: Fraction of Explained Variance ρ and Measure of Sampling Adequacy

		$T = 10$		$T = 20$		$T = 30$	
Method	Statistic	$\beta = .2$	$\beta = .8$	$\beta = .2$	$\beta = .8$	$\beta = .2$	$\beta = .8$
Untransformed	ρ	.398	.562	.247	.363	.200	.297
	MSA	.051	.859	.108	.930	.132	.948
Limited (Ltd.)	ρ	.350	.457	.197	.279	.154	.224
	MSA	.028	.776	.079	.901	.112	.931
Collapsed (Col.)	ρ	.700	.911	.670	.917	.669	.923
	MSA	.938	.999	.974	1.000	.981	1.000
Ltd. & Col.	ρ	.828	.968	.766	.966	.748	.967
	MSA	.926	.999	.977	1.000	.987	1.000

The explained variance from PCA of the collapsed (and limited) instrument set is in the high .70s, low .80s for $\beta = .2$ and in the high .90s for $\beta = .8$. Almost all of the variation of the standard instrument set can be explained by much fewer components. Irrespective of β, PCAs of the untransformed or limited instrument set do not score appreciable values (cf. Figure 4.5). This is the main reason why these procedures fail to result in plausible estimates (cf. Table 4.1). Although high

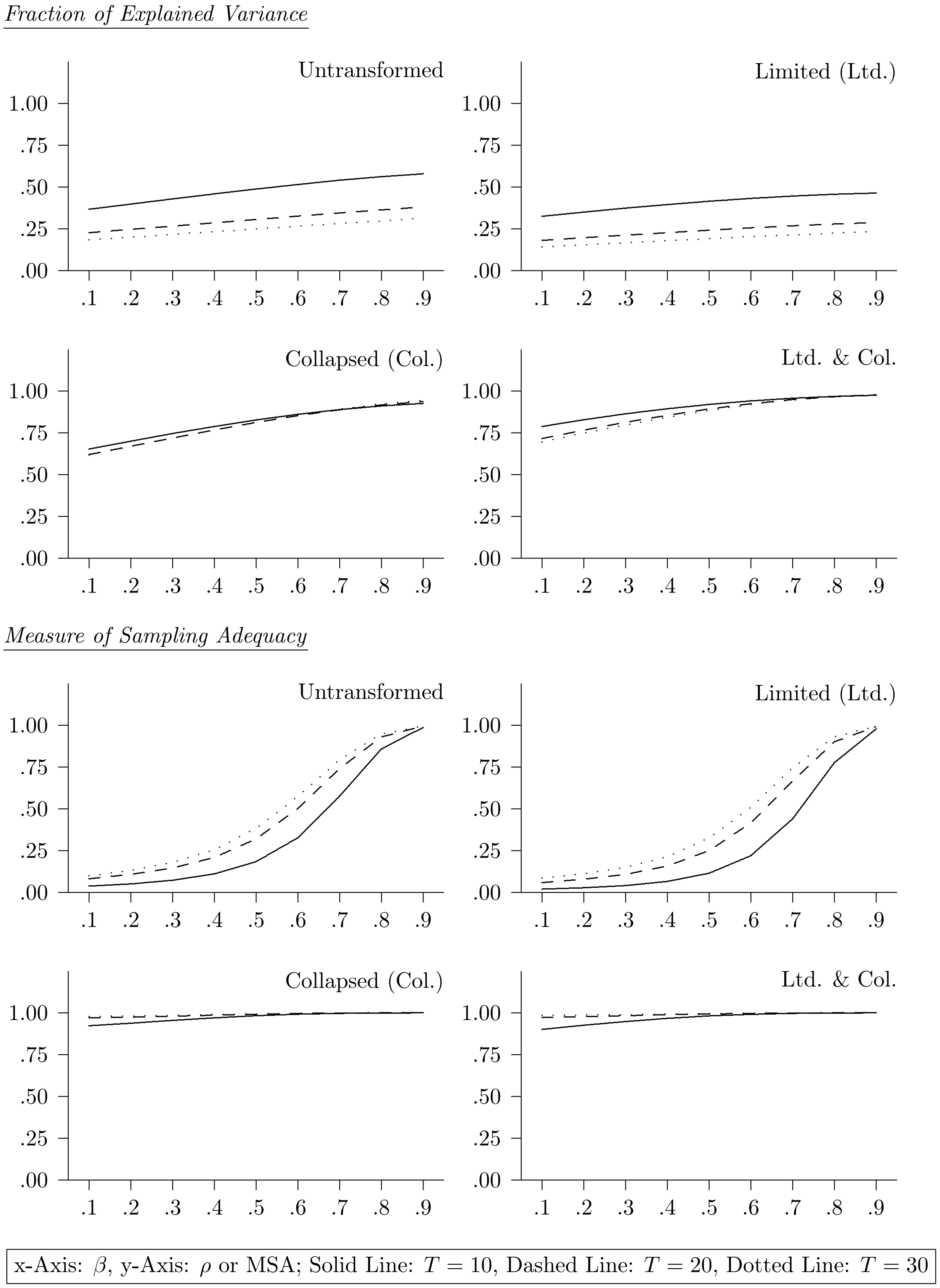


Figure 4.5: Fraction of Explained Variance ρ and Measure of Sampling Adequacy

MSAs can be achieved for $\beta = .8$, the explained variance remains low. MSAs for the first two procedures are close to one in all instances. The collapsed instrument set is much more suitable for PCA as each instrument is non-zero for all applicable observations, unlike untransformed instruments which are non-zero for just a single observation.

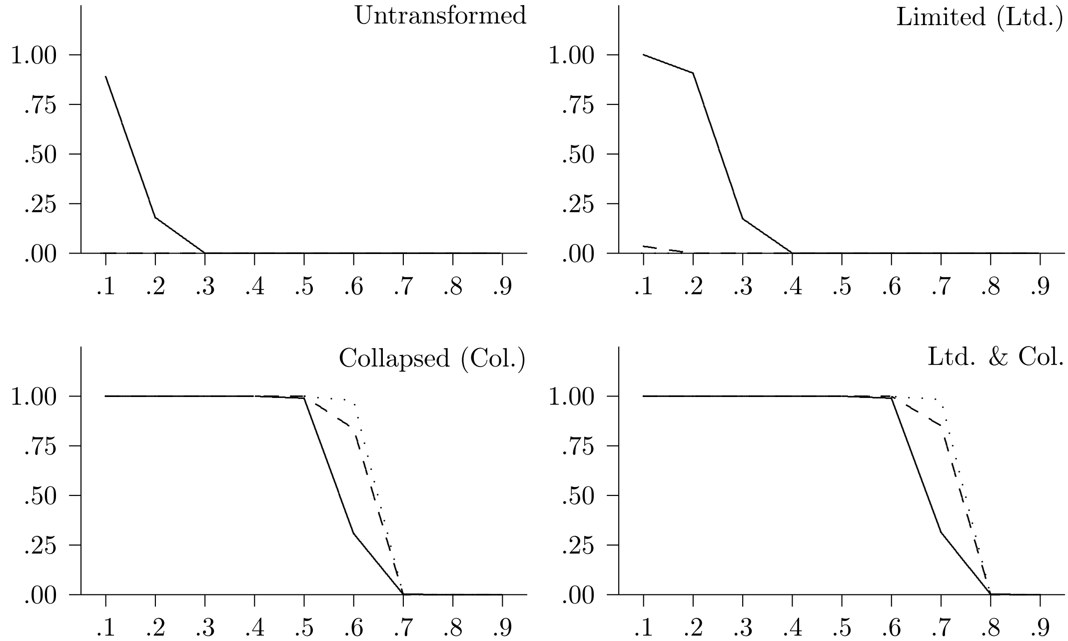
Table 4.4 gives the proportions for which the weakness of the instruments has been rejected, along with the number of instruments employed in the estimation for each of the methods used. As a rule of thumb, the instrument set is deemed to be weak if the F-statistic from the first stage regression in two stages least squares (TSLS) is less than ten. This is an approximate test at the 5% significance level that the TSLS bias is at most 10% of the OLS bias (Stock and Yogo, 2005).

Table 4.4: Instrument Count J and Rejection Frequency of Weak Instruments

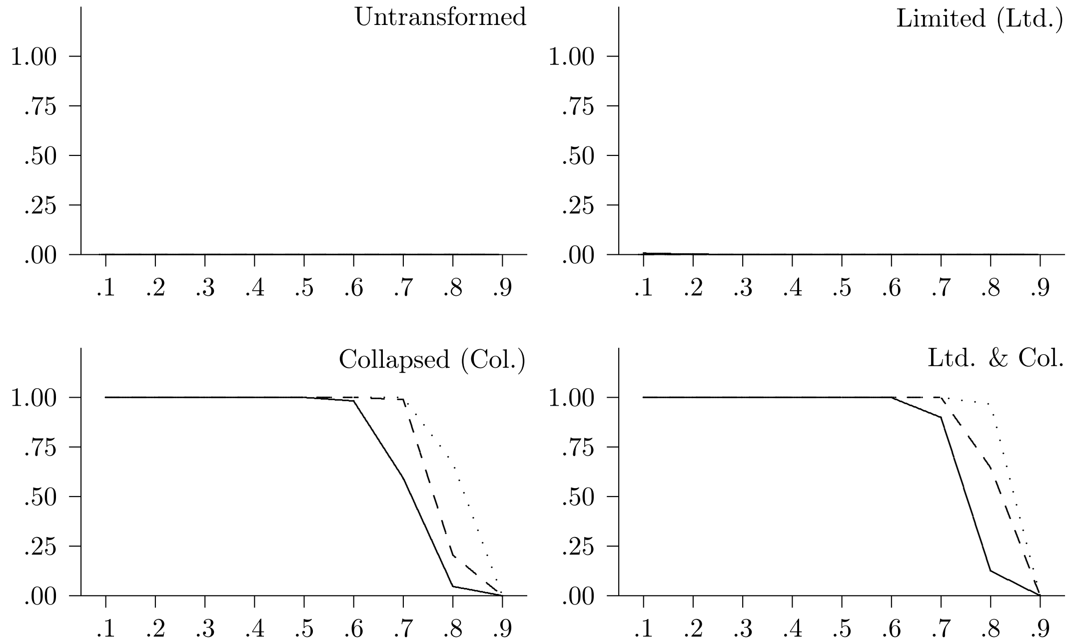
	$T = 10$			$T = 20$			$T = 30$		
Method	J	$\beta = .2$	$\beta = .8$	J	$\beta = .2$	$\beta = .8$	J	$\beta = .2$	$\beta = .8$
Standard GMM									
Untransformed	36	.180	.000	171	.000	.000	406	.000	.000
Limited (Ltd.)	26	.908	.000	126	.000	.000	301	.000	.000
Collapsed (Col.)	8	1.000	.000	18	1.000	.000	28	1.000	.000
Ltd. & Col.	4	1.000	.002	9	1.000	.000	14	1.000	.000
Factorised GMM									
Untransformed	3	.000	.000	4	.000	.000	5	.000	.000
Limited (Ltd.)	3	.002	.000	4	.000	.000	5	.000	.000
Collapsed (Col.)	2	1.000	.047	3	1.000	.205	4	1.000	.670
Ltd. & Col.	2	1.000	.125	3	1.000	.646	4	1.000	.966

Irrespective of the instrument set used, the instruments get weaker, the higher β becomes (cf. Figure 4.6). This is because the more the process approaches a random walk, the lower is the correlation between levels and differences. Both in standard and factorised GMM, as T rises, the untransformed or limited instrument set becomes weaker, while the collapsed (and limited) instrument set gets stronger. Partial R^2s of deeper, uncollapsed instruments are virtually zero; thus, these add almost no new information for estimation. Moreover, it seems as if many weak instruments cause the entire instrument set to be weak even though it contains a few strong ones. Again, the factorised variants of the collapsed (and limited) instrument set perform best, while the factorised variants of the untransformed

Standard GMM



Factorised GMM



x-Axis: β, y-Axis: f; Solid Line: $T = 10$, Dashed Line: $T = 20$, Dotted Line: $T = 30$

Figure 4.6: Rejection Frequency f of Weak Instruments

or limited instrument set are worse than their standard GMM counterparts. Factorised instruments are the only ones which are strong even for relatively high β and any T.

4.4 Directions for Applied Research

The Monte Carlo results strongly suggest the use of factorised instruments as these produce the lowest bias and RMSE. This generates an ultimate set of instruments and reduces the uncertainty researchers face in their choice of instruments. Furthermore, there is a clear recommendation to collapse the instrument set prior to factorisation or, if factorisation is not to be used at all, then at the very least the instrument set should be collapsed. To reiterate, this implies a deterministic transformation of the standard instrument set, and the factorised variant of this instrument set is the method of choice. Preferably, the lag depth is also limited. The lag limit should be chosen based on a priori information on the value of the autoregressive parameter, as serial correlation decreases exponentially. Most importantly, standard GMM suffers from instrument proliferation. The findings in this chapter indicate that results of numerous applications of GMM in the literature may benefit from factorised instruments. LSDV should be applied only if the time dimension is much larger than 30, while pooled OLS should not be used at all in the estimation of DPD.

In applied research, the number of retained components from PCA can be derived from factor analytic criteria, such as MSA, and should be tested for their validity in the GMM framework. The methodology outlined here can be applied to System GMM or exogenous variables in a completely analogous fashion. It is reasonable to make use of the correlation between all instruments to lower the instrument count.

Appendix A: Proof of Theorems

Proof of Theorem 4.1. Using the definition of $\mathbf{Z}^*$ in Theorem 4.1 and Equation (4.2), the proposition follows directly from the linearity property of the expectation operator: $\mathrm{E}(\mathbf{Z}^{*\prime}\boldsymbol{\Delta\varepsilon}) = \mathrm{E}(\mathbf{F}'\mathbf{Z}'\boldsymbol{\Delta\varepsilon}) = \mathbf{F}'\mathrm{E}(\mathbf{Z}'\boldsymbol{\Delta\varepsilon}) = \mathbf{0}$. □

Proof of Theorem 4.2. Per definitionem of Theorem 4.2, the corresponding elements of $\mathbf{Z}$ and $\boldsymbol{\Delta\varepsilon}$, meaning those which form the cross products in $\mathbf{Z}'\boldsymbol{\Delta\varepsilon}$, are independent random variables, and thus Borel. For any pair $\phi(\cdot)$ and $\psi(\cdot)$ of Borel functions, the corresponding elements of $\phi(\mathbf{Z})$ and $\psi(\boldsymbol{\Delta\varepsilon})$ are also independent.

$\widehat{\mathrm{Var}}(\mathbf{Z})$ is a positive semi-definite symmetric matrix meaning that all eigenvalues are real and non-negative. It is well-established that the sum and product of two real-valued measurable functions are measurable. That eigenvectors can be found in a Borel measurable fashion was shown by Azoff (1974, Corollary 4).

Hence, the corresponding elements of $\mathbf{Z}^{**} = \mathbf{Z}\Lambda(\mathbf{Z}) = \phi(\mathbf{Z})$, with $\mathbf{F}^* = \Lambda(\mathbf{Z})$ being the matrix of component loadings, and $\boldsymbol{\Delta\varepsilon} = \psi(\boldsymbol{\Delta\varepsilon})$ are independent random variables, too. Moreover, given quadratic integrability of the elements of $\mathbf{Z}^{**}$ and $\boldsymbol{\Delta\varepsilon}$, the corresponding ones are uncorrelated. The proposition follows from the fact that this can be the case if and only if $\mathrm{E}(\mathbf{Z}^{**\prime}\boldsymbol{\Delta\varepsilon}) = \mathbf{0}$ as $\mathrm{E}(\boldsymbol{\Delta\varepsilon}) = \mathbf{0}$. □

Appendix B: Structure of Transformation Matrices

For the sake of exposition, let $T = 6$ and $i = 1, 2, \ldots, N$. Note that the first observation is dropped due to differencing.

Untransformed

The standard instrument set consists of lagged values of the endogenous variable; in particular, one instrument is generated for each time period and lag available.

$$\frac{1}{N}\sum_{i=1}^{N} y_{i,t-1-\ell}\Delta\hat{\varepsilon}_{i,t} = 0\,\forall\, t = 3,\, 4\, \ldots,\, T \wedge \ell = 1,\, 2,\, \ldots,\, t-2 \tag{4.3}$$

$$\mathbf{Z_i} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ y_{i,1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & y_{i,2} & y_{i,1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & y_{i,3} & y_{i,2} & y_{i,1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & y_{i,4} & y_{i,3} & y_{i,2} & y_{i,1} \end{bmatrix}$$

The instrument count is $J = (T-2)(T-1)/2 = 10$ for standard instruments and $\tilde{J} = \left\lceil \sqrt[4]{J} \right\rceil$ for factorised instruments, where $\lceil \cdot \rceil$ is the ceiling function.

Limited (L)

Limiting the maximum lag depth of $y_{i,t-1}$ to $\tau = 2$, in general $\tau = (T-2)/2$, gives as transformation matrix a block matrix of identity matrices up to dimension τ (for each time period, indicated by solid lines) separated by rows of zeroes (for excluded lags, indicated by dashed lines).

$$\frac{1}{N} \sum_{i=1}^{N} y_{i,t-1-\ell} \Delta \hat{\varepsilon}_{i,t} = 0 \,\forall\, t = 3,\, 4\, \ldots,\, T \wedge \ell = 1,\, 2,\, \ldots,\, \tau,\, \tau \leq t-2 \tag{4.4}$$

$$\mathbf{Z_i^L} = \mathbf{Z_i F^L} = \mathbf{Z_i} \left[\begin{array}{c|cc|cc|cc} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hdashline 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right] = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ y_{i,1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & y_{i,2} & y_{i,1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & y_{i,3} & y_{i,2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & y_{i,4} & y_{i,3} \end{bmatrix}$$

The instrument count becomes $J^L = J - (T-2-\tau)(T-1-\tau)/2 = 7$ and $\tilde{J}^L = \left\lceil \sqrt[4]{J^L} \right\rceil$, respectively.

Collapsed (C)

The transformation matrix for collapsing the instrument set is made up of identity matrices of increasing dimension stacked one upon the other (indicated by solid lines) with blocks of zero matrices to the right (indicated by dashed lines).

$$\frac{1}{N(T-2)}\sum_{i=1}^{N}\sum_{t=3}^{T} y_{i,t-1-\ell}\Delta\hat{\varepsilon}_{i,t} = 0\,\forall\,\ell = 1,\,2,\,\ldots,\,t-2 \tag{4.5}$$

$$\mathbf{Z_i^C} = \mathbf{Z_i F^C} = \mathbf{Z_i}\left[\begin{array}{c:ccc} 1 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ y_{i,1} & 0 & 0 & 0 \\ y_{i,2} & y_{i,1} & 0 & 0 \\ y_{i,3} & y_{i,2} & y_{i,1} & 0 \\ y_{i,4} & y_{i,3} & y_{i,2} & y_{i,1} \end{bmatrix}$$

By collapsing the instrument count is cut to $J^C = T-2 = 4$ and $\tilde{J}^C = \left\lceil \sqrt[3]{J^C} \right\rceil$, respectively.

Limited & Collapsed (LC)

When both techniques are combined, i.e. rows of zeroes from $\mathbf{F^L}$ and stacked identity matrices (now again only up to dimension τ) from $\mathbf{F^C}$.

$$\frac{1}{N(T-2)}\sum_{i=1}^{N}\sum_{t=3}^{T} y_{i,t-1-\ell}\Delta\hat{\varepsilon}_{i,t} = 0\,\forall\,\ell = 1,\,2,\,\ldots,\,\tau,\,\tau \le t-2 \tag{4.6}$$

$$\mathbf{Z_i^{LC}} = \mathbf{Z_i F^{LC}} = \mathbf{Z_i} \left[\begin{array}{c:c} 1 & 0 \\ \hline 1 & 0 \\ 0 & 1 \\ \hline 1 & 0 \\ 0 & 1 \\ \hdashline 0 & 0 \\ \hline 1 & 0 \\ 0 & 1 \\ \hdashline 0 & 0 \\ 0 & 0 \end{array}\right] = \begin{bmatrix} 0 & 0 \\ y_{i,1} & 0 \\ y_{i,2} & y_{i,1} \\ y_{i,3} & y_{i,2} \\ y_{i,4} & y_{i,3} \end{bmatrix}$$

Using this technique reduces the instrument count to $J^{LC} = \tau = 2$ and $\tilde{J}^{LC} = \left\lceil \sqrt{J^{LC}} \right\rceil$, respectively.

Appendix C: GMM Estimator and Test Statistics

Generalised Method of Moments

Following Hansen (1982), GMM is an asymptotically efficient estimation method if instruments outnumber regressors, i.e. if the specification is overidentified. The task is to estimate the parameter vector $\boldsymbol{\beta}$, given data matrices $\mathbf{Y}$ and $\mathbf{X}$ in a linear model of the form:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{E}. \tag{4.7}$$

The regressors are thought to be endogenous, i.e. correlated with the error term $\mathbf{E}$. Hence, instrumental variables $\mathbf{Z}$ are needed which are orthogonal to the the latter. Assume that the researcher has these instruments $\mathbf{Z}$ at hand, so that:

$$\mathrm{E}(\mathbf{Z}'\mathbf{E}) = \mathbf{0} \text{ in addition to } \mathrm{E}(\mathbf{E}|\mathbf{Z}) = \mathbf{0}. \tag{4.8}$$

If instruments outnumber regressors, the corresponding vector of empirical moments will not be exactly zero generally. To overcome this problem, GMM minimises a generalised metric of the empirical moments of the errors with the instruments, based on a full-rank matrix $\mathbf{A}$ that weights moments:

$$\hat{\boldsymbol{\beta}}_{\mathbf{A}} = \arg\min_{\hat{\boldsymbol{\beta}}} \left\| \frac{1}{n} \mathbf{Z}'\hat{\mathbf{E}} \right\|_{\mathbf{A}}, \tag{4.9}$$

where the empirical residuals $\hat{\mathbf{E}}$ are

$$\hat{\mathbf{E}} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}. \tag{4.10}$$

This metric takes the following positive semi-definite quadratic form:

$$\left\| \frac{1}{n} \mathbf{Z}'\hat{\mathbf{E}} \right\|_{\mathbf{A}} = n \left(\frac{1}{n} \mathbf{Z}'\hat{\mathbf{E}} \right)' \mathbf{A} \frac{1}{n} \mathbf{Z}'\hat{\mathbf{E}} = \frac{1}{n} \hat{\mathbf{E}}'\mathbf{Z}\mathbf{A}\mathbf{Z}'\hat{\mathbf{E}}. \tag{4.11}$$

Minimisation requires the first matrix derivative with respect to the estimated parameter vector $\hat{\boldsymbol{\beta}}$ to be zero, where the last step uses Equations (4.11) and (4.10), respectively:

$$\frac{\partial}{\partial \hat{\boldsymbol{\beta}}} \left\| \frac{1}{n} \mathbf{Z}'\hat{\mathbf{E}} \right\|_{\mathbf{A}} = \frac{\partial}{\partial \hat{\mathbf{E}}} \left\| \frac{1}{n} \mathbf{Z}'\hat{\mathbf{E}} \right\|_{\mathbf{A}} \frac{\partial}{\partial \hat{\boldsymbol{\beta}}} \hat{\mathbf{E}} = -\frac{2}{n} \hat{\mathbf{E}}'\mathbf{Z}\mathbf{A}\mathbf{Z}'\mathbf{X} \stackrel{!}{=} \mathbf{0}. \tag{4.12}$$

This expression is solved for the linear GMM estimator by first transposing it, second substituting Equation (4.10) and eventually inverting the leading term:

$$\mathbf{X}'\mathbf{Z}\mathbf{A}\mathbf{Z}'\mathbf{X}\hat{\boldsymbol{\beta}}_{\mathbf{A}} = \mathbf{X}'\mathbf{Z}\mathbf{A}\mathbf{Z}'\mathbf{Y} \Longleftrightarrow \hat{\boldsymbol{\beta}}_{\mathbf{A}} = (\mathbf{X}'\mathbf{Z}\mathbf{A}\mathbf{Z}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}\mathbf{A}\mathbf{Z}'\mathbf{Y}. \tag{4.13}$$

Efficient GMM requires that moments are weighted in inverse proportion to their (co-)variances. Hence, the weighting matrix is the inverse of the variance matrix of the moments:

$$\mathbf{A} = \mathrm{Var}(\mathbf{Z}'\mathbf{E})^{-1} = (\mathbf{Z}'\mathrm{Var}(\mathbf{E}|\mathbf{Z})\mathbf{Z})^{-1} = (\mathbf{Z}'\boldsymbol{\Omega}\mathbf{Z})^{-1}. \tag{4.14}$$

DPD Model

In the dynamic panel data framework of Arellano and Bond (1991) the above data matrices are stacked by cross-section and are defined as follows: $\mathbf{Y} = \boldsymbol{\Delta}\mathbf{y} = \mathbf{My}$, $\mathbf{X} = \boldsymbol{\Delta}\mathbf{y_{-1}} = \mathbf{My_{-1}}$ and $\mathbf{E} = \boldsymbol{\Delta\varepsilon} = \mathbf{M}\boldsymbol{\varepsilon}$, where $\mathbf{M} = \mathbf{I_N} \otimes \mathbf{M_i}$ with $\mathbf{M_i}$ being a $[(T-2) \times (T-1)]$ matrix of the form:

$$\mathbf{M_i} = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & \cdots \\ 0 & -1 & 1 & 0 & 0 & \cdots \\ 0 & 0 & -1 & 1 & 0 & \cdots \\ 0 & 0 & 0 & -1 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix}.$$

The structure of the instrument matrix $\mathbf{Z}$ can be inferred from Appendix B. The parameter vector collapses to the single lagged endogenous variable parameter: $\boldsymbol{\beta} = \beta$. The number of observations for Difference GMM is: $n = N(T-2)$.

However, $\boldsymbol{\Omega}$ is in general unknown and needs to be estimated. A first stage proxy, using the above definitions and the assumption that the untransformed errors are i.i.d., is:

$$\boldsymbol{\Omega} = \mathrm{Var}(\mathbf{M}\boldsymbol{\varepsilon}|\mathbf{Z}) = \mathbf{M}\mathrm{Var}(\boldsymbol{\varepsilon}|\mathbf{Z})\mathbf{M}' \xrightarrow{\mathrm{Var}(\boldsymbol{\varepsilon}|\mathbf{Z})=\sigma^2\mathbf{I}} \sigma^2\mathbf{MM}' = \sigma^2\mathbf{H}, \tag{4.15}$$

where $\mathbf{H} = \mathbf{I_N} \otimes \mathbf{H_i}$ with $\mathbf{H_i} = \mathbf{M_i}\mathbf{M_i}'$ being a quadratic $[(T-2)]$ matrix of the form:

$$\mathbf{H_i} = \begin{bmatrix} 2 & -1 & 0 & 0 & \cdots \\ -1 & 2 & -1 & 0 & \cdots \\ 0 & -1 & 2 & -1 & \cdots \\ 0 & 0 & -1 & 2 & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix}.$$

Thus, the feasible DPD GMM estimator is:

$$\hat{\beta} = (\boldsymbol{\Delta}\mathbf{y_{-1}}'\mathbf{Z}(\mathbf{Z}'\mathbf{HZ})^{-1}\mathbf{Z}'\boldsymbol{\Delta}\mathbf{y_{-1}})^{-1}\boldsymbol{\Delta}\mathbf{y_{-1}}'\mathbf{Z}(\mathbf{Z}'\mathbf{HZ})^{-1}\mathbf{Z}'\boldsymbol{\Delta}\mathbf{y}. \tag{4.16}$$

Monte Carlo Simulation

The results of this chapter are based on $B = 1,000$ replications. The pseudocode for the simulation reads:

Pseudocode	Input	Output
Simulate data	$\boldsymbol{\alpha}$, β, $\boldsymbol{\varepsilon}$, $\mathbf{y_0}$, N, T	$\mathbf{y}$
Generate instrument set	$\mathbf{y_{-1-\ell}}$	$\mathbf{Z}$ ($\mathbf{Z}^*$)
Factorise instruments	$\mathbf{Z}$ ($\mathbf{Z}^*$)	$\mathbf{Z}^{**}$, ρ, MSA
Estimate parameter	$\mathbf{\Delta y}$, $\mathbf{\Delta y_{-1}}$, $\mathbf{Z}^{**}$, $\mathbf{H}$	$\hat{\beta}$, S, F

Bias, root mean squared error and standard deviation are defined as follows:

$$\text{Bias} = \frac{1}{B}\sum_{b=1}^{B}\hat{\beta}_b - \beta = \bar{\hat{\beta}} - \beta, \tag{4.17}$$

$$\text{RMSE} = \sqrt{\frac{1}{B}\sum_{b=1}^{B}\left(\hat{\beta}_b - \beta\right)^2} \text{ and} \tag{4.18}$$

$$\text{SD} = \sqrt{\frac{1}{B}\sum_{b=1}^{B}\left(\hat{\beta}_b - \bar{\hat{\beta}}\right)^2}. \tag{4.19}$$

Note the obvious relation between the three, that is:

$$\text{RMSE}^2 = \text{SD}^2 + \text{Bias}^2. \tag{4.20}$$

The Sargan (1958) test statistic S for the validity of overidentifying restrictions is a Wald test with the null hypothesis of joint validity, i.e. it tests whether or not the vector of empirical moments is randomly distributed around $\mathbf{0}$:

$$S = \left(\frac{1}{n}\mathbf{Z}'\mathbf{\Delta}\hat{\boldsymbol{\varepsilon}}\right)' \text{Var}\left(\frac{1}{n}\mathbf{Z}'\mathbf{\Delta}\hat{\boldsymbol{\varepsilon}}\right)^{\mathbf{-1}} \frac{1}{n}\mathbf{Z}'\mathbf{\Delta}\hat{\boldsymbol{\varepsilon}} = \frac{1}{n\hat{\sigma}^2}\mathbf{\Delta}\hat{\boldsymbol{\varepsilon}}'\mathbf{Z}(\mathbf{Z}'\mathbf{H}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{\Delta}\hat{\boldsymbol{\varepsilon}}. \tag{4.21}$$

Explained variance ρ from PCA is the proportion of the first f eigenvalues of the estimated covariance matrix of the instruments to all J eigenvalues:

$$\rho = \frac{\sum_{j=1}^{f} \lambda_j}{\sum_{j=1}^{J} \lambda_j}. \tag{4.22}$$

The Kaiser-Meyer-Olkin (Kaiser, 1970) measure of sampling adequacy (MSA) depends on the so-called anti-image of the correlation matrix. (For the sake of simplicity, the presentation focuses on correlations rather than on covariances.) Let $\mathbf{R^{-1}}$ be the inverse of the correlation matrix $\mathbf{R}$, where the former should be near diagonal for "sampling adequacy". Then $\mathbf{P} = \mathbf{SR^{-1}S}$ is the anti-image of the covariance matrix, where $\mathbf{S} = (\operatorname{diag}\mathbf{R^{-1}})^{-\mathbf{1}}$ with $\operatorname{diag}\mathbf{R^{-1}}$ consisting of the diagonal elements of $\mathbf{R^{-1}}$ only. The anti-image of the correlation matrix is then defined as $\mathbf{Q} = \mathbf{TPT}$, where $\mathbf{T} = (\operatorname{diag}\mathbf{P})^{-\mathbf{1/2}}$. The off-diagonal elements of $\mathbf{Q}$, $q_{k,l}$, $k \neq l$, are the negatives of the partial correlations. MSA follows as the ratio of the sum of squared correlations to the sum of squared correlations plus the sum of squared partial correlations:

$$\text{MSA} = \frac{\sum_{k=1}^{J-1} \sum_{l=k+1}^{J} r_{k,l}^2}{\sum_{k=1}^{J-1} \sum_{l=k+1}^{J} r_{k,l}^2 + \sum_{k=1}^{J-1} \sum_{l=k+1}^{J} q_{k,l}^2}. \tag{4.23}$$

MSA is obviously bounded to lie between zero and one; the higher it is, the better. Kaiser and Rice (1974) label the outcome as follows: $.9 \leq \text{MSA} \leq 1.0$ "marvellous", $.8 \leq \text{MSA} < .9$ "meritorious", $.7 \leq \text{MSA} < .8$ "middling", $.6 \leq \text{MSA} < .7$ "mediocre", $.5 \leq \text{MSA} < .6$ "miserable" and $.0 \leq \text{MSA} < .5$ "unacceptable".

The Staiger and Stock (1997), and Stock and Yogo (2005) F-statistic for the test of weak instruments under the null hypothesis stems from the first stage regression in TSLS of the endogenous variable on the instruments and is a function of the R^2:

$$F = \frac{(n-J)R^2}{J(1-R^2)}. \tag{4.24}$$

Chapter 5

Summary

This thesis takes the views of both users and producers of official statistics. At three stages – processing, validation and analysis of data – the problem of data adequacy is discussed. In empirical research, the degree of uncertainty is an important issue. Hence, to be able to make an informed decision it is crucial that one looks closely not only into the results but also into the data and methodology used. Otherwise, the interpretation might be misleading.

Processing of data is analysed in the first main chapter. From the perspective of producers of price indices, the aggregation of price data without having information on quantities or expenditures is a common problem. The chapter's contribution to the literature is the proposition of a statistical approach that allows the achievement of numerical equivalence between an elementary index and a desired aggregate index based on the price elasticity alone. In an empirical application data from German foreign trade statistics is analysed. The results indicate that a whole range of elementary indices would be necessary in the calculation of price indices in order to mirror the fluctuations of the aggregate index. However, the findings suggest a pronounced preference for the Carli index at the elementary level of a Laspeyres price index. This translates into the harmonic index if the Paasche index should be approximated.

The second main chapter deals with validation of data. Producers as well as users of seasonally adjusted data are interested in the quality and interpretation of business cycle indicators. Revisions to seasonally adjusted real time data stem from two sources, revisions from the seasonal adjustment method and those from unadjusted real time data. A new procedure for decomposition of revisions is developed and contributes to the literature. Five important German business cycle indicators are considered in the empirical application. Predictability of time series and orders of magnitude of their components are made responsible for the revision properties of the time series. The decomposition reveals that revisions of unadjusted real time data, implying the use of newly available information, play a larger role than those from the seasonal adjustment method, which are of purely technical nature.

In the third main chapter data analysis is examined from the user perspective. Estimation of econometric models with dynamic panel data suffers either from DPD bias (LSDV) or instrument proliferation (GMM). As a solution to the latter problem, a new methodology for reducing the instrument count in instrumental variables estimation is developed and is a contribution to the literature. Factorisation of the standard instrument set is shown to be a valid and data-driven transformation, i.e. it results in a consistent estimator whose instruments are stochastically transformed. Applied to the limited and/or collapsed instrument set, this results in the lowest bias and RMSE in a Monte Carlo study. Furthermore, the overidentifying restrictions are more robust as regards both the exogeneity and the relevance assumption. Notably, the estimates of standard GMM are heavily downward biased and the instruments are very poor in small samples.

Bibliography

Allen, R.G.D. (1975), *Index Numbers in Theory and Practice*, London, United Kingdom: Macmillan.

Anderson, T.W., and Hsiao, C. (1982), "Formulation and Estimation of Dynamic Models Using Panel Data," *Journal of Econometrics*, 18, 47-82.

Arellano, M., and Bond, S. (1991), "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58, 277-297.

Arellano, M., and Bover, O. (1995), "Another Look at the Instrumental Variable Estimation of Error-Components Models," *Journal of Econometrics*, 68, 29-51.

Azoff, E.A. (1974), "Borel Measurability in Linear Algebra," *Proceedings of the American Mathematical Society*, 42, 346-350.

Balk, B.M. (1994), "On the First Step in the Calculation of a Consumer Price Index," in: Ducharme, L.M. (Ed.), *Proceedings of the First Meeting of the International Working Group on Price Indices*, Ottawa, Canada: Statistics Canada.

Balk, B.M. (1998), "On the Use of Unit Value Indices as Consumer Price Subindices," in: Lane, W. (Ed.), *Proceedings of the Fourth Meeting of the International Working Group on Price Indices*, Washington, DC: Bureau of Labor Statistics.

Balk, B.M. (2005), "Price Indexes for Elementary Aggregates: The Sampling Approach," *Journal of Official Statistics*, 21, 675-699.

Balk, B.M. (2008), *Price and Quantity Index Numbers*, New York, NY: Cambridge University Press.

Blundell, R., and Bond, S. (1998), "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87, 115-143.

von Bortkiewicz, L. (1923), "Zweck und Struktur einer Preisindexzahl", *Nordisk Statistisk Tidskrift*, 2, 369-408.

Boskin, M.J., Dulberger, E.R., and Griliches, Z. (1996), *Toward a More Accurate Measure of the Cost of Living*, Darby, PA: Diane Publishing.

Boskin, M.J., Dulberger, E.R., Gordon, R.J., Griliches, Z., and Jorgenson, D.W. (1998), "Consumer Prices, the Consumer Price Index, and the Cost of Living," *Journal of Economic Perspectives*, 12, 3-26.

Breitung, J. (2000), "The Local Power of Some Unit Root Tests for Panel Data," in: Baltagi, B.H. (Ed.), *Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, in: *Advances in Econometrics*, 15, Amsterdam, The Netherlands: JAI Press.

Carli, G.R. (1764), "Del Valore e della Proporzione de' Metalli Monetati con i Generi in Italia Prima delle Scoperte dell' Indie col Confronto del Valore e della Proporzione de' Tempi Nostri," in: *Opere Scelte di Carli, Vol. 1*, Milan, Italy.

Carruthers, A.G., Sellwood, D.J., and Ward, P.W. (1980), "Recent Developments in the Retail Prices Index," *The Statistician*, 29, 1-32.

Coggeshall, F. (1887), "The Arithmetic, Geometric and Harmonic Means," *Quarterly Journal of Economics*, 1, 83-86.

Dalén, J. (1992), "Computing Elementary Aggregates in the Swedish Consumer Price Index," *Journal of Official Statistics*, 8, 129-147.

Dalén, J. (1995), "Quantifying Errors in the Swedish Consumer Price Index," *Journal of Official Statistics*, 11, 261-275.

Dalén, J. (1998), "Studies on the Comparability of Consumer Price Indices," *International Statistical Review*, 66, 83-113.

Dalén, J. (1999), "A Note on the Variance of the Sample Geometric Mean," *Department of Statistics, Stockholm University Research Report*, 1.

Damia, V., and Picón Aguilar, C. (2006), "Quantitative Quality Indicators for Statistics – An Application to Euro Area Balance of Payment Statistics," *ECB Occasional Paper Series*, 54.

Deaton, A. (1985), "Panel Data from Time Series of Cross-Sections," *Journal of Econometrics*, 30, 109-126.

Deutsche Bundesbank (1999), "The Changeover from the Seasonal Adjustment Method Census X-11 to Census X-12-ARIMA," *Deutsche Bundesbank Monthly Report*, September, 39-50.

Dickey, D.A., and Fuller, W.A. (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, 74, 427-431.

Diewert, W.E. (1976), "Exact and Superlative Index Numbers," *Journal of Econometrics*, 4, 115-145.

Diewert, W.E. (1978), "Superlative Index Numbers and Consistency in Aggregation," *Econometrica*, 46, 883-900.

Diewert, W.E. (1995), "Axiomatic and Economic Approaches to Elementary Price Indexes," *NBER Working Paper*, 5104.

Diewert, W.E. (2004), "Elementary Indices," in: ILO, IMF, OECD, UNECE, Eurostat, and World Bank, *Consumer Price Index Manual: Theory and Practice*, ed. by Hill, P., Geneva, Switzerland: International Labour Organization.

Diewert, W.E., and Silver, M. (2004), "Elementary Indices," in: IMF, ILO, OECD, UNECE, and World Bank, *Producer Price Index Manual: Theory and Practice*, ed. by Armknecht, P., Washington, DC: International Monetary Fund.

Diewert, W.E., and Silver, M. (2008), "Elementary Indices," in: IMF, ILO, OECD, UNECE, and World Bank, *Export and Import Price Index Manual*, ed. by Silver, M., http://www.imf.org/external/np/sta/tegeipi/index.htm, 10 April 2008.

Drobisch, M.W. (1871), "Ueber die Berechnung der Veränderungen der Waarenpreise und des Geldwerths," *Jahrbücher für Nationalökonomie und Statistik*, 16, 143-156.

Dutot, C. (1738), *Réflexions politiques sur les finances et le commerce*, The Hague, The Netherlands: Les frères Vaillant & N. Prevost.

Eichhorn, W. (1978), *Functional Equations in Economics*, Reading, MA: Addison-Wesley.

Eichhorn, W., and Voeller, J. (1976), *Theory of the Price Index: Fisher's Test Approach and Generalizations*, in: *Lecture Notes in Economics and Mathematical Systems*, 140, Berlin, Germany: Springer.

European Commission (1996), "Commission Regulation (EC) No 1749/96," *Official Journal of the European Communities*, L 229, 3-10.

European Commission (2001), *Compendium of HICP Reference Documents*, Luxembourg: Office for Official Publications of the European Communities.

European Commission (2006), *Statistics on the Trading of Goods: User Guide*, Luxembourg: Office for Official Publications of the European Communities.

Eurostat (2006), *Proceedings of the Conference on Seasonality, Seasonal Adjustment and their Implications for Short-Term Analysis and Forecasting*, Luxembourg, 10-12 May 2006, Luxembourg: Office for Official Publications of the European Communities.

Eurostat (2009), *ESS Guidelines on Seasonal Adjustment*, Luxembourg: Office for Official Publications of the European Communities.

Feenstra, R.C., and Shapiro, M.D. (Eds.) (2003), *Scanner Data and Price Indexes*, Proceedings of the Conference on Research in Income and Wealth, Arlington, VA, 15-16 September 2000, Chicago, IL: University of Chicago Press.

Fenwick, D. (1999), "The Impact of Choice of Base Month and Other Factors on the Relative Performance of Different Formulae Used for Aggregation of Consumer Price Index Data at an Elementary Aggregate Level," in: Guðnason, R., and Gylfadóttir, Þ. (Eds.), *Proceedings of the Fifth Meeting of the International Working Group on Price Indices*, Reykjavik, Iceland: Statistics Iceland.

Findley, D., Monsell, B., Bell, W., Otto, M., and Chen, B.-C. (1998), "New Capabilities and Methods of the X-12-ARIMA Adjustment Program," *Journal of Business and Economic Statistics*, 16, 127-152.

Fisher, I. (1922), *The Making of Index Numbers*, Boston, MA: Houghton Mifflin.

Fisher, R.A. (1925), *Statistical Methods for Research Workers*, Edinburgh, United Kingdom: Oliver & Boyd.

Gerberding, C., Seitz, F., and Worms, A. (2005), "How the Bundesbank Really Conducted Monetary Policy," *North American Journal of Economics and Finance*, 16, 277-292.

Hansen, L.P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029-1054.

Hill, R.J. (2006), "Superlative Index Numbers: Not All of Them Are Super," *Journal of Econometrics*, 130, 25-43.

Holtz-Eatkin, D., Newey, W., and Rosen, H.S. (1988), "Estimating Vector Autoregressions with Panel Data," *Econometrica*, 56, 1371-1395.

Im, K.S., Pesaran, M.H., and Shin, Y. (2003), "Testing for Unit Roots in Heterogeneous Panels," *Journal of Econometrics*, 115, 53-74.

Jevons, W.S. (1863), *A Serious Fall in the Value of Gold Ascertained, and its Social Effects Set Forth*, London, United Kingdom: Stanfords.

Jevons, W.S. (1865), "On the Variation of Prices and the Value of the Currency since 1782," *Journal of the Statistical Society of London*, 28, 294-325.

Judson, R.A., and Owen, A.L. (1999), "Estimating Dynamic Panel Data Models: A Guide for Macroeconomists," *Economics Letters*, 65, 9-15.

Kaiser, H.F. (1970), "A Second Generation Little Jiffy," *Psychometrika*, 35, 401-415.

Kaiser, H.F., and Rice, J. (1974), "Little Jiffy, Mark IV," *Educational and Psychological Measurement*, 34, 111-117.

Kirchner, R. (1999), "Implications of the New Seasonal Adjustment Method Census X-12-ARIMA for Current Economic Analysis in Germany," *Research Centre of the Deutsche Bundesbank Discussion Paper*, 7.

Kiviet, J.F. (1995), "On Bias, Inconsistency, and Efficiency of Various Estimators in Dynamic Panel Data Models," *Journal of Econometrics*, 68, 53-78.

Knetsch, T., and Reimers, H.-E. (2009), "Dealing with Benchmark Revisions in Real-Time Data: The Case of German Production and Orders Statistics," *Oxford Bulletin of Economics and Statistics*, 71, 209-235.

Konüs, A.A. (1924), "The Problem of the True Index of the Cost of Living," *Economic Bulletin of the Institute of Economic Conjuncture* (in Russian), 9-10, 64-71; *Econometrica*, 7 (1939), 10-29.

Ladiray, D., and Quenneville, B. (2001), *Seasonal Adjustment with the X-11 Method*, *Lecture Notes in Statistics*, 158, New York, NY: Springer.

Laspeyres, E. (1871), "Die Berechnung einer mittleren Waarenpreissteigerung," *Jahrbücher für Nationalökonomie und Statistik*, 16, 296-314.

Levin, A., Lin, C.F., and Chu, C.S.J. (2002), "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties," *Journal of Econometrics*, 108, 1-24.

von der Lippe, P. (2001), *Chain Indices: A Study in Price Index Theory*, Stuttgart, Germany: Metzler-Poeschel, in: *Spectrum of Federal Statistics*, 16, Wiesbaden, Germany: Federal Statistical Office.

von der Lippe, P. (2007), *Index Theory and Price Statistics*, Frankfurt am Main, Germany: P. Lang.

von der Lippe, P., and Mehrhoff, J. (2009), "Unit Value Bias Reconsidered," in: Becker Vermeulen, C. (Ed.), *Proceedings of the Eleventh Meeting of the International Working Group on Price Indices*, Neuchâtel, Switzerland: Swiss Federal Statistical Office.

Lippi, M., and Forni, M. (1990), "On the Dynamic Specification of Aggregated Models," in: Barker, T., and Pesaran, M.H. (Eds.), *Disaggregation in Econometric Modelling*, London, United Kingdom: Routledge.

Lloyd, P.J. (1975), "Substitution Effects and Biases in Nontrue Price Indices," *American Economic Review*, 65, 301-313.

Mankiw, N.G., and Shapiro, M.D. (1986), "News or Noise? An Analysis of GNP Revisions," *NBER Working Paper*, 1939.

Mehrhoff, J. (2007), "A Linear Approximation to the Jevons Index," in: von der Lippe, P., *Index Theory and Price Statistics*, Frankfurt am Main, Germany: P. Lang.

Moulton, B.R. (1993), "Basic Components of the CPI: Estimation of Price Changes," *Monthly Labor Review*, December.

Moulton, B.R. (1996), "Constant Elasticity Cost-of-Living Index in Share Relative Form," mimeo, Washington, DC: Bureau of Labor Statistics.

Newey, W.K., and West, K.D. (1994), "Automatic Lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 61, 631-653.

Nickel, S. (1981), "Biases in Dynamic Models with Fixed Effects," *Econometrica*, 49, 1417-1426.

OECD (2007), *Proceedings of the Meeting of the OECD Short-term Economic Statistics Working Party – OECD/Eurostat Taskforce on Performing Revisions Analysis for Sub-Annual Economic Statistics*, Paris, 25-27 June 2007, Paris, France: OECD Publishing.

Orphanides, A. (2001), "Monetary Policy Rules Based on Real Time Data," *American Economic Review*, 91, 964-985.

Paasche, H. (1874), "Ueber die Preisentwicklung der letzten Jahre nach den Hamburger Börsennotirungen," *Jahrbücher für Nationalökonomie und Statistik*, 23, 168-178.

Párniczky, G. (1974), "Some Problems of Price Measurement in External Trade Statistics", *Acta Oeconomica*, 12, 229-240.

Phillips, P.C.B., and Perron, P. (1988), "Testing for a Unit Root in Time Series Regression," *Biometrika*, 75, 335-346.

Roodman, D. (2009), "A Note on the Theme of Too Many Instruments," *Oxford Bulletin of Economics and Statistics*, 71, 135-158.

Sargan, J.D. (1958), "The Estimation of Economic Relationships Using Instrumental Variables," *Econometrica*, 26, 393-415.

Schultze, C.L., and Mackie, C. (Eds.) (2002), *At What Price? Conceptualizing and Measuring Cost-of-Living and Price Indexes*, Washington, DC: National Academy Press.

Selvanathan, E.A., and Prasada Rao, D.S. (2002), *Index Numbers: A Stochastic Approach*, London, United Kingdom: Macmillan.

Shapiro, M.D., and Wilcox, D.W. (1997), "Alternative Strategies for Aggregating Prices in the CPI," *Federal Reserve Bank of St. Louis Review*, May/June, 113-125.

Silver, M. (1995), "Elementary Aggregates, Micro-Indices, and Scanner Data: Some Issues in the Compilation of Consumer Price Indices," *Review of Income and Wealth*, 41, 427-438.

Silver, M. (2009), "Do Unit Value Export, Import, and Terms-of-Trade Indices Misrepresent Price Indices?," *IMF Staff Papers*, 56, 297-322.

Silver, M., and Heravi, S. (2007), "Why Elementary Price Index Number Formulas Differ: Evidence on Price Dispersion," *Journal of Econometrics*, 140, 874-883.

Silver, M., and Webb, B. (2002), "The Measurement of Inflation: Aggregation at the Basic Level," *Journal of Economic and Social Measurement*, 28, 21-35.

Staiger, D., and Stock, J.H. (1997), "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65, 557-586.

Stock, J.H., and Yogo, M. (2005), "Testing for Weak Instruments in Linear IV Regression," in: Andrews, D.W.K., and Stock, J.H. (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, New York, NY: Cambridge University Press.

Szulc, B.J. (1989), "Price Indices Below the Basic Aggregation Level," in: Turvey, R., *Consumer Price Indices*, Geneva, Switzerland: International Labour Office.

Szulc, B.J. (1994), "Choice of Price Index Formulae at the Micro-Aggregation Level: The Canadian Empirical Evidence," in: Ducharme, L.M. (Ed.), *Proceedings of the First Meeting of the International Working Group on Price Indices*, Ottawa, Canada: Statistics Canada.

Törnqvist, L. (1936), "The Bank of Finland's Consumption Price Index," *Bank of Finland Monthly Bulletin*, 10, 1-8.

Turvey, R. (1996), "Elementary Aggregate (Micro) Indexes," in: European Commission, *Improving the Quality of Price Indices: CPI and PPP*, Proceedings of the Eurostat International Seminar, Florence, 18-20 December 1995, Luxembourg: Office for Official Publications of the European Communities.

U. S. Census Bureau (2001), *X-12-ARIMA Reference Manual*, Version 0.2.8.

Walsh, C.M. (1901), *The Measurement of General Exchange-Value*, New York, NY: Macmillan.

Walsh, C.M. (1921), *The Problem of Estimation*, London, United Kingdom: P.S. King & Son.

Windmeijer, F. (2005), "A Finite Sample Correction for the Variance of Linear Efficient Two-Step GMM Estimators," *Journal of Econometrics*, 126, 25-51.